SPATIOTEMPORAL DYNAMICS OF NEURAL REINSTATEMENT DURING PAIRED ASSOCIATES MEMORY

by
Robert B. Yaffe

A dissertation submitted to the Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
March, 2016

1 Abstract

Reinstatement of neural activity is hypothesized to underlie our ability to mentally travel back in time in order to recover the context of a previous experience. We used intracranial recordings to directly examine the precise spatiotemporal extent of neural reinstatement as 32 participants with electrodes placed for seizure monitoring performed a paired associates episodic verbal memory task. By cueing recall, we were able to compare reinstatement during correct and incorrect trials, and found that successful retrieval occurs with reinstatement of a gradually changing neural signal present during encoding. We examined reinstatement in individual frequency bands and individual electrodes and found that neural reinstatement was largely mediated by temporal lobe theta and high gamma frequencies. Leveraging the high temporal precision afforded by intracranial recordings, our data demonstrate that high gamma activity associated with reinstatement preceded theta activity during encoding, but during retrieval this difference in timing between frequency bands was absent. Using a time-warping algorithm, we quantify the time scale of this replay and show that on average, retrieval exhibits replay of the spectral dynamics of encoding on a faster timescale. This supports theories that memory retrieval consists of a temporally compressed replay of encoding. Our results complement previous studies showing that distributed patterns of neural activity are coordinated to encode a representation of temporal context which is then recovered during successful memory retrieval with precise spatiotemporal dynamics.

Dissertation Readers:

Sridevi Sarma, Ph.D.

Kareem Zaghloul, MD, Ph.D.

2 Acknowledgments

There are countless people who I need to thank who have helped and inspired me along the way. I want to start by thanking Matthew Kerr because he is the one who recruited me to the Sarma Lab at Hopkins. He took me under his wing and I am very grateful for all the help, guidance, and friendship he gave me throughout grad school. I would like to thank my advisors, Dr. Sri Sarma and Dr. Kareem Zaghoul for all they have taught me. I had a unique opportunity to do my graduate work at both Hopkins and the NIH and this was possible thanks to Sri and Kareem. They are both excellent teachers and have inspired me to pursue a career in academia. I would also like to thank John Wittig Jr., Zar Zavala, Julio Chapeton, Rafi Haque, Srikanth Damera, Andrew Yang, Alex Vaz, Timothy Sheehan, Jennifer Arai, Sam Burns, Sabato Santaniello, Kevin Kahn, Rahul Agarwal, Duluxan Sritharan, Austin Jordan, and Yaqing Su for their invaluable help along the way. I would like to thank my committee members Dr. Stan Anderson and Dr. Nathan Crone for their support.

A very special thank you to my grandparents Dr. Arthur and Sandy Becker who made it possible for me to work at NIH by taking me in as a boarder and adding me to their meal plan. Thank you to my parents Carrie and Gordon who have always been my biggest fans and have supported me every step of the way. Thanks to my brother, Neal, for always being there for me. Thank you to my aunts, uncles, and cousins who have shared various parts of the journey with me. Thanks to my roommate throughout all of grad school, Mel, it has been an honor living in the Mushroom Palace with you. And to Kari, thank you for inspiring me to always live on the edge of adventure.

Contents

3 Introduction

Reinstatement of neural activity is hypothesized to underlie our ability to recover the internal representation of a previous experience, a process described as mental time travel [2, 34, 38, 49]. These internal representations, which may reflect the external environment or internal mental states, form the context in which an episodic memory is embedded. Central to the hypothesis of mental time travel is that context representations in the brain gradually change over time, and that successful retrieval of an episodic memory involves mentally jumping back in time to re-experience a particular context [18, 19, 41, 45]. Consistent with this paradigm, when an episode is successfully retrieved from memory, the memory for neighboring episodes that occurred close in time is enhanced, an effect known as contiguity [28]. Yet despite substantial behavioral data supporting this hypothesis, a number of important yet unanswered questions remain regarding its underlying neural mechanisms.

Empiric support for neural reinstatement has largely emerged from functional MRI studies that have used multi-voxel pattern analysis (MVPA) [7, 25, 42]. MVPA relies on classifying neural activity during retrieval in order to dissociate broad manipulations such as category or task that are present during encoding [24, 29, 35, 40]. MVPA, however, is unable to directly examine whether successful retrieval reinstates the neural representations of individual items. Representational similarity analysis supports neural reinstatement of individual stimuli, [10, 43, 47], but this alternative fMRI approach does not examine to what extent retrieval reinstates a changing neural representation of context. In addition, the limited temporal resolution of fMRI studies make them unable to identify the precise spatiotemporal dynamics of neural activity that distinguish successful and unsuccessful retrieval of individual events.

Intracranial EEG (iEEG) recordings offer an opportunity to explore the neural mechanisms of reinstatement with high temporal and spatial precision. Spectral power during retrieval has recently been shown to resemble activity present dur-

ing encoding of neighboring items, providing neural evidence for the contiguity effect particular to mental time travel [32]. However, because retrieval in this study was unconstrained, one possibility is that the observed reinstatement was a consequence, rather than a cause, of retrieval contiguity. Spiking activity in the medial temporal lobe (MTL) has also been shown to reinstate activity present during encoding [20,36]. But it is unknown how the temporal dynamics of neural activity distinguish reinstatement between correct and incorrect memory retrieval across broader cortical regions.

We directly investigated whether reinstatement of a gradually changing neural representation of context occurs during successful retrieval and explore the precise spatiotemporal dynamics that underlie this process. We investigate the reinstatement of spectral power across multiple frequency bands and multiple electrode locations during encoding and retrieval as patients with implanted subdural electrodes participated in a verbal paired associates memory task. Paired associates memory tasks offer an opportunity to directly explore these questions as explicit associations are formed, and subsequently recalled, between pairs of items [17,27]. Hence, rather than rely on a participant's own retrieval strategy, we use direct experimental control over retrieval to examine the relationship between retrieval cues and the reinstatement of a drifting context representation.

Additionally, many studies have shown that increased synchronous brain activity is a biomarker of successful memory encoding [26,50]. It remains unknown whether there are specific patterns of brain connectivity that are present during encoding which are reinstated during retrieval. In this study, we identify functional connectivity networks using a mutual information (MI) metric and show that patterns of connectivity are reinstated between encoding and retrieval.

A complete understanding of the spatiotemporal dynamics of the neural mechanisms of memory formation and retrieval will be of great value in the development of brain stimulation paradigms targeted to enhance memory function. Previous ef-

forts to enhance memory function through electrical stimulation [48] have focused on targets such as the hippocampus and entorhinal cortex, which have been shown to be crucial for memory formation. However, many cortical areas are involved in the formation and retrieval of memories [17, 22, 46]. By studying the precise electrophysiological signatures of memory throughout the entire cortex, other cortical regions may be implicated as stimulation targets. Furthermore, the relationship between stimulation parameters and the behavioral effects of stimulation are unknown. With a clear understanding of the dynamics of spectral power in different frequency bands and the underlying connectivity, the parameters of stimulation could also be targeted to induce activity that is conducive to successful memory encoding and retrieval.

4 Intracranial EEG (iEEG) recordings

In our study, participants with medication-resistant epilepsy underwent a surgical procedure in which platinum recording contacts were implanted subdurally on the cortical surface as well as deep within the brain parenchyma. For each participant, the clinical team determined the placement of the contacts so as to best localize epileptogenic regions. Data were collected at three different hospitals: the Clinical Center at the National Institutes of Health (NIH; Bethesda, MD), the Hospital of the University of Pennsylvania (UP; Philadelphia, PA), and Thomas Jefferson University Hospital (TJ; Philadelphia, PA). The research protocol was approved by the institutional review board at each hospital, and informed consent was obtained from the participants and their guardians.

Whereas each hospital used the same general implantation procedures and data-acquisition techniques, our analysis had to account for technical details that varied by institution. Intracranial EEG (iEEG) data were recorded using a Nihon Kohden (NIH and TJ) or Nicolet (UP) EEG data acquisition system. Depending on the amplifier and the discretion of the clinical team, EEG signals were sampled at 512, 1000, or 2000 Hz. Signals were referenced to a common contact placed subcutaneously, on the scalp, or on the mastoid process. All recorded traces were resampled at 1000 Hz, and a fourth order 2 Hz stopband butterworth notch filter was applied at 60 Hz to eliminate electrical line noise. The testing laptop sent +/-5 V analog pulses via an optical isolator into a pair of open lines on the clinical recording system to synchronize the electrophysiological recordings with behavioral events.

We collected electrophysiological data from a total of 2670 subdural and depth recording contacts (PMT Corporation, Chanhassen, MN; AdTech, Racine, WI). Subdural contacts were arranged in both grid and strip configurations with an inter-contact spacing of 10 mm. Hippocampal depth electrodes (6-8 linearly arranged contacts) were placed in three patients. Contact localization was accomplished by co-

registering the post-op CTs with the post-op MRIs using both FSL Brain Extraction Tool (BET) and FLIRT software packages and mapped to both MNI and Talairach space using an indirect stereotactic technique and OsiriX Imaging Software DICOM viewer package. The resulting contact locations were subsequently projected to the cortical surface of a Montreal Neurological Institute N27 standard brain [12]. Pre-op MRI's were used when post-op MR images were were not available.

We analyzed iEEG data using bipolar referencing in order to reduce volume conduction and confounding interactions between adjacent electrodes [39]. We defined the bipolar montage in our data-set based on the geometry of iEEG electrode arrangements. For every grid, strip, and depth probe, we isolated all pairs of contacts that were positioned immediately adjacent to one another; bipolar signals were then found by differencing the signals between each pair of immediately adjacent contacts. The resulting bipolar signals were treated as new virtual electrodes (referred to as electrodes throughout the text), originating from the mid-point between each contact pair. All subsequent analyses were performed using these derived bipolar signals. In total, our dataset consisted of 2779 electrodes (1154 left hemispheric, 1625 right hemispheric).

5 Encoding and Retrieval

5.1 Insights from Free Recall Task

To understand the interactions between encoding and retrieval, we must first understand the mechanisms of each process individually. Previous studies by Burke, et al. [4, 5] have analyzed iEEG during encoding and retrieval of a free recall task. In this task, a list of words is presented sequentially during the encoding period. Next, there is a delay period with a distractor task. Finally, the subject is asked to recall as many words as possible from the list in any order.

Analyzing the data from this task, it was shown that high gamma (64 to 95Hz) power increased during successful memory encoding and retrieval. During encoding, two networks were identified that showed significantly greater high gamma power during the encoding of words that were subsequently recalled compared to words that were forgotten. The first network showed early activation, and is believed to be involved in visual processing of the word. The second network showed subsequent activation, and is believed to be involved in the memory storage of the word. The early activation network is composed of the occipital lobes and the inferior temporal lobes. The late activation network is composed of the left prefrontal cortex, and the left posterior temporal lobe.

During memory retrieval, three networks were identified. The right temporal lobe showed increased theta power during the onset of retrieval. Then, the medial temporal lobes, the left prefrontal cortex, and the left posterior temporal lobe showed increased high gamma power during memory retrieval. Lastly, the bilateral motor language areas showed increases in high gamma power with vocalization of the recalled word. Note that the regions activated during the late encoding period believed to be responsible for memory storage are also activated during the memory retrieval

period. This provides some evidence of reinstatement of general encoding and retrieval mechanisms during free recall. In our study, we looked for these general encoding and retrieval mechanisms during a paired associates memory task.

5.2 Paired Associates Task

32 participants (18 male; age 33.5 $\pm$ 2.2 years) with medication resistant epilepsy who underwent a surgical procedure for placement of intracranial electrodes for seizure monitoring participated in a verbal paired associates task (Fig. 1). Lists were composed of four pairs of common nouns, chosen at random and without replacement from a pool of high-frequency nouns. Words were presented sequentially and appeared in capital letters at the center of the screen. Study word pairs were separated from their corresponding recall cue by a minimum lag of two study or test items. During the study period (encoding), each word pair was preceded by an orientation stimulus (a row of capital X's) that appeared on the screen for 300 ms followed by a blank interstimulus interval (ISI) of 750 ms with a jitter of 75 ms. Word pairs were then presented on the screen for 2500 ms followed by a blank ISI of 1500 ms with a jitter of 75 ms. During the test period (retrieval), one randomly chosen word from each study pair was shown, and the participant was asked to recall the other word from the pair by vocalizing a response into a microphone. Each cue word was preceded by an orientation stimulus (a row of question marks) that appeared on the screen for 300 ms followed by a blank ISI of 750 ms with a 75 ms jitter. Cue words were then presented on the screen for 3000 ms followed by a blank ISI of 4500 ms. Participants could vocalize their response any time during the recall period after cue presentation. Vocalizations were digitally recorded and then manually scored for analysis. Responses were designated as correct, as intrusions, or as passes when no vocalization was made or when the participant vocalized the word 'pass'. Intrusion and pass trials were designated as incorrect trials. A single experimental session contained up to 25 lists.

For analysis, we designated a single trial as the encoding period for a study word pair and the retrieval period during testing of its corresponding cue.

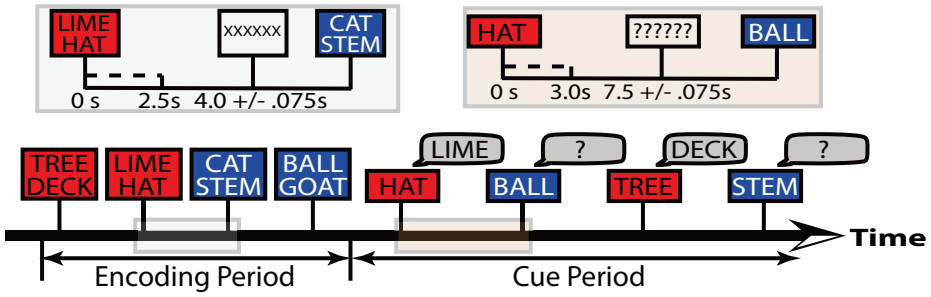

Figure 1 Paired associates task presentation. Timing of word and cue presentation is shown in the inset. Red and blue boxes indicate correct and incorrect trials, respectively.

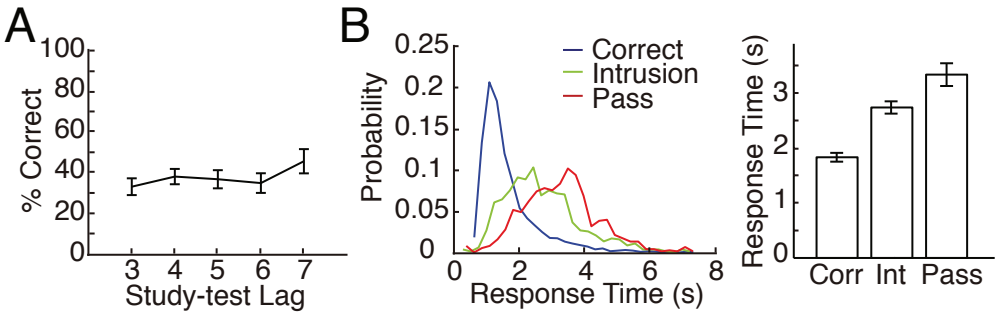

Figure 2 (*A*) Probability of correct recall for all study-test lags averaged across all participants. Error bars represent standard error of the mean (SEM). (*B*) Response time distributions (*left*) and average mean response times (*right*) for correct, intrusion, and pass trials across all participants. Error bars represent SEM.

Participants studied 223 ± 23 word pairs and successfully recalled 41.0 ± 8.0% words with a mean response time of 1831 ± 80 ms (Fig. 2B). On 11.8 ± 1.6% of trials, participants responded with an incorrect word (intrusions). The mean response time for intrusions was 2736 ± 116 ms. For the remaining study word pairs, participants either made no response to the cue word, or vocalized the word 'pass'. Participants vocalized the word 'pass' with a mean response time of 3334 ± 207 ms. A one-way ANOVA revealed no significant effect of study-test lag on recall probability ($F(4, 155) = 1.06$, $P = 0.38$) (Fig. 2A).

To identify the general mechanisms that are involved in successful encoding and

retrieval, we analyzed the progressions of theta and high gamma power. To do this, we created brain plots showing the average z-scored power across subjects for each frequency band for each 500ms window during the task. Figure 3 shows the progression of activity during encoding. The regions that show increases in high gamma during successful encoding are similar to the regions that Burke, et al. [4] found to show increases in high gamma during free recall encoding. This suggests a common memory encoding mechanism across free recall and paired associates. This is not unexpected because both tasks involve the visual presentation of a written word or words that are supposed to be memorized. However, it is not a trivial finding because there are substantial differences between the tasks. To summarize the findings, paired associates encoding is characterized by an early increase in high gamma power in the occipital lobes and the inferior temporal lobes and an early decrease in high gamma power in the frontal lobes. This is followed by an increase in high gamma in the left posterior temporal lobe and the left prefrontal cortex. The theta power generally shows the same effects but in the opposite direction. High gamma and theta tend to be negatively correlated [4].

Figure 4 shows the progression of activity during successful memory retrieval. Here, we see that similarly to encoding, there is an early high gamma increase in the occipital lobes and inferior temporal lobes. This is believed to be a neural signature of visual processing of a word or words on the screen. As retrieval progresses, there is a high gamma increase in the medial temporal lobes and the prefrontal cortices. This pattern of activation is similar to the pattern of activation seen in spontaneous retrieval during free recall [5]. Here, we have an explicit retrieval cue that begins the retrieval process. It appears from our data that there is a progression of activity from the occipital lobes to the inferior temporal lobes and to the medial temporal lobes which causes the memory retrieval. The similarities between the progressions of activity during encoding and retrieval should also be noted. During both encoding

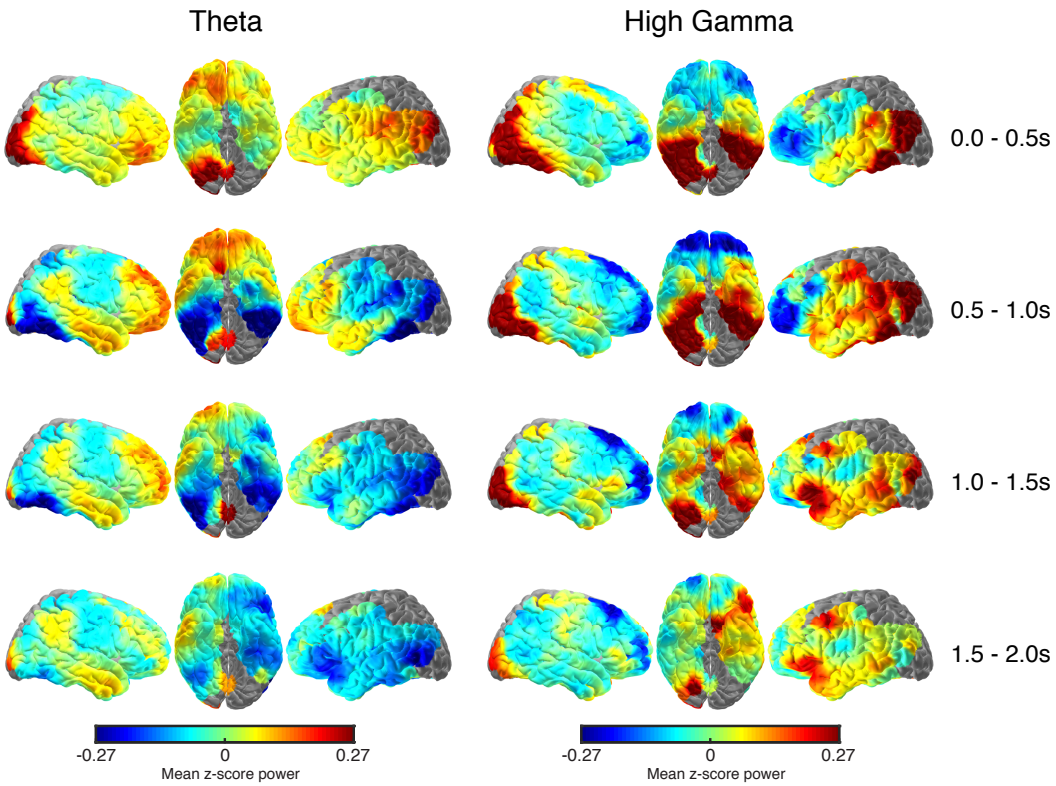

Figure 3 The mean z-scored theta (3.5 - 8Hz) and high gamma (62-100Hz) power is shown during four 500ms intervals during encoding. Time 0 is when the pair is shown on the screen. Data is shown for successful encoding events averaged across subjects.

and retrieval, there is activity related to the processing of a visual cue and activity related to the memory encoding or retrieval of information. There is considerable overlap between the regions that are involved in encoding and those involved in retrieval. Lastly during retrieval, we see activation of bilateral speech motor areas that are involved in the vocalization of a response. Note that the average time of vocalization for a correct trial was 1831ms. Again we observe that the patterns of theta activity generally inversely mirror the high gamma activity patterns. One additional observation here is that the high gamma patterns of activation appears to be more spatially precise than the broad activity patterns of theta.

Now that we have explored the general mechanisms of paired associates encoding and retrieval, we will now consider the interactions between the two. The main aim of this work is to quantify the relationship between encoding and retrieval.

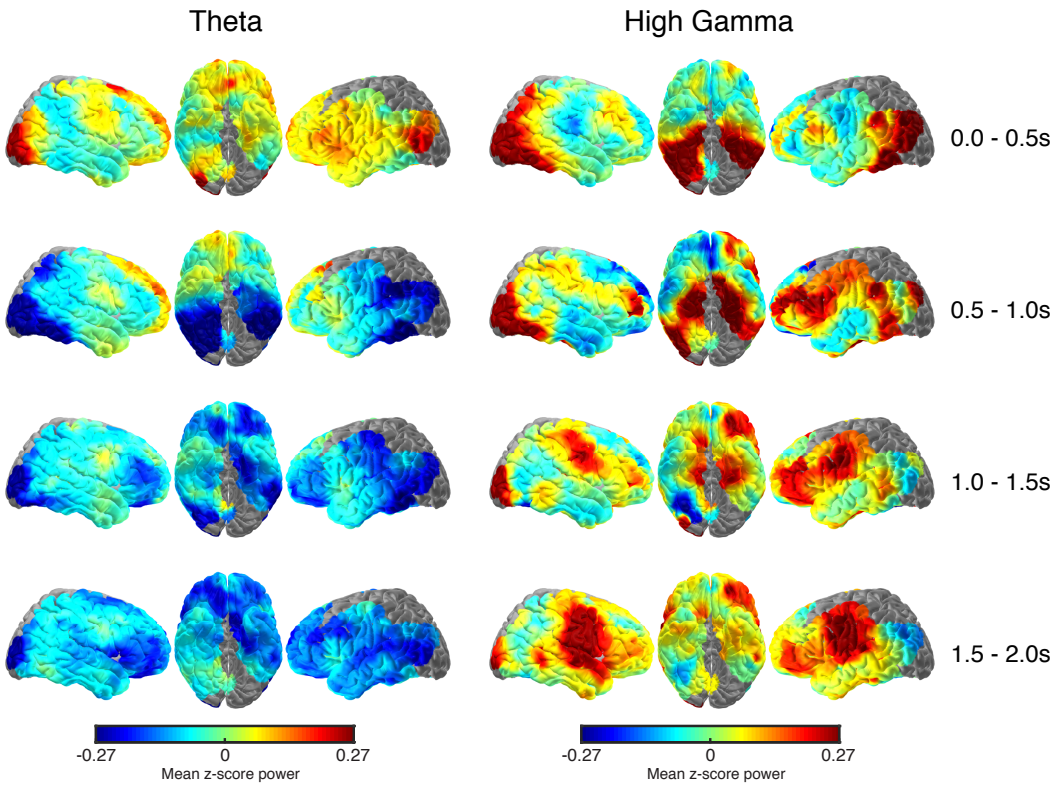

Figure 4 The mean z-scored theta (3.5 - 8Hz) and high gamma (62-100Hz) power is shown during four 500ms intervals during retrieval. Time 0 is when the cue is shown on the screen. Data is shown for correct retrieval events averaged across subjects. Mean response time is 1831ms.

6 Reinstatement of Spectral Power

6.1 Quantifying Reinstatement of Spectral Power

To assess the reinstatement of neural activity between encoding and retrieval, we compared multi-dimensional representations of spectral power at every time point during encoding and retrieval. Briefly, for every 500 ms temporal epoch, spaced every 100 ms (80% overlap), during the encoding and retrieval periods, we constructed a feature vector containing spectral power information from five frequency bands (theta 3.5 - 8 Hz; alpha 8 - 12 Hz; beta 13 - 25 Hz; low gamma 30 - 58 Hz; high gamma 62 - 100 Hz) and from all electrode locations. We quantified reinstatement between every

temporal epoch during encoding and retrieval by calculating the cosine similarity between feature vectors. Thus, for a single trial, we generate a precise temporal map of neural reinstatement between the encoding and retrieval periods (Fig. 5).

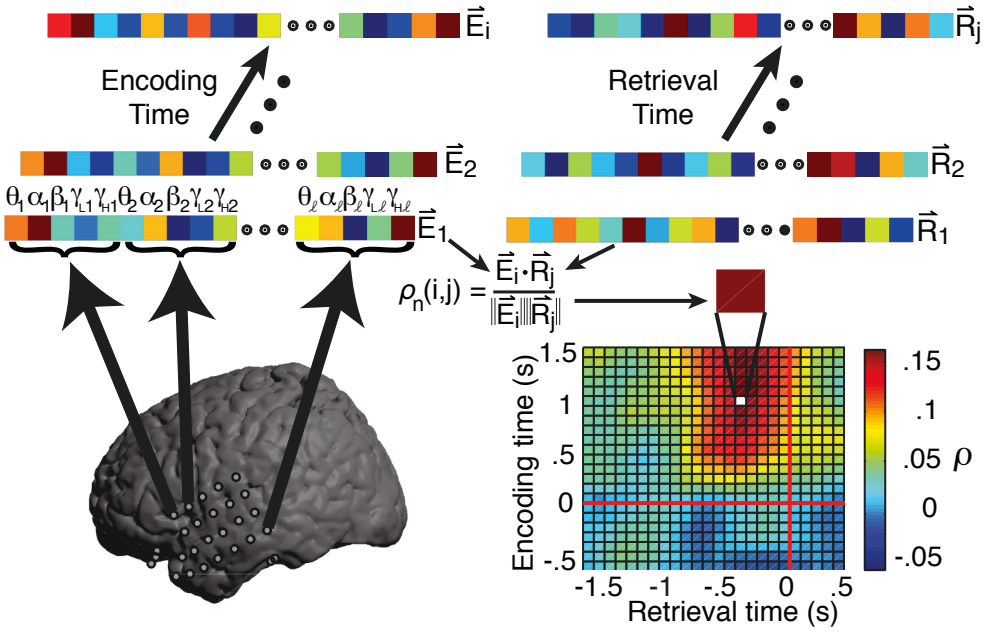

Figure 5 Reinstatement of distributed spectral power. For every temporal epoch during encoding and retrieval, the z-scored power from every electrode and every frequency band is combined to create a single feature vector for that epoch. Reinstatement (cosine similarity) for all encoding-retrieval pairs is shown for a single trial. The red lines correspond to the time of word presentation during encoding and the time of vocalization during retrieval.

To compute the spectral power, we used 46 logarithmically-spaced wavelets ranging from 2-100 Hz. During the encoding period, we convolved each wavelet with 6000 ms of iEEG data surrounding the presentation of word pairs, from 1500 ms before stimulus presentation to 4500 ms after stimulus presentation. During the retrieval period, we convolved each wavelet with 9500 ms of iEEG data. Data were time-locked to vocalization and extended from 8000 ms before vocalization to 1500 ms after vocalization. For trials in which patients did not respond, vocalization response time was defined as a time randomly drawn from the distribution of response times during the correct trials for that patient. These response times were not included in the later analysis of response time distributions. One participant was instructed

to vocalize her response only after the cue word disappeared from the screen. Her response times were also not included in analysis of response time distributions. In all cases, we included a 1000 ms buffer on both sides of the clipped data.

To quantify changes in spectral power during encoding and retrieval, we convolved the bipolar iEEG signals with complex valued Morlet wavelets (wavelet number 6) to obtain magnitude and phase information [1]. We squared and log-transformed the magnitude of the continuous-time wavelet transform to generate a continuous measure of instantaneous power.

We binned the continuous time transforms into 500 ms epochs spaced every 100 ms (80% overlap) and averaged the instantaneous power over each epoch. Each 500 ms epoch is marked with the time that corresponds to the end of the temporal epoch. These time markings are used to indicate the time on the axes of the reinstatement maps and in the calculations of peak reinstatement time.

To account for changes in power across experimental sessions, we z-transformed power values separately for each frequency and for each session using the mean and standard deviation of all 500 ms epochs for that session. For each temporal epoch, we subsequently averaged the z-transformed power across five frequency bands: theta (3.5Hz to 8Hz), alpha (8Hz to 12Hz), beta (13Hz to 25Hz), low gamma (30Hz to 58Hz), and high gamma (62 Hz to 100Hz).

For every temporal epoch in each trial, we constructed a feature vector comprised of the average z-scored power for every electrode and for every frequency band. For each encoding temporal epoch, i, and for each retrieval temporal epoch, j, we define feature vectors as:

$$\vec{E}_i = [z_{1,1}(i) \ \ldots \ z_{1,F}(i) \ \ldots \ z_{L,F}(i)]$$

$$\vec{R}_j = [z_{1,1}(j) \ \ldots \ z_{1,F}(j) \ \ldots \ z_{L,F}(j)]$$

where $z_{l,f}(i)$ is the z-transformed power of electrode $l = 1 \ldots L$ at frequency band $f = 1 \ldots F$ in temporal epoch i. For L electrodes and F frequency bands, we thus create a feature vector at each temporal epoch that contains $K = L * F$ features.

To quantify reinstatement during trial n, we calculated the cosine similarity between all encoding and retrieval feature vectors $\vec{E}_i$ and $\vec{R}_j$ for all pairs of encoding and retrieval temporal epochs during that trial (Fig. 5). Cosine similarity gives a measure of how close the angles of two vectors are in a multidimensional space. We chose cosine similarity over Pearson's correlation to measure reinstatement because if all the elements of two feature vectors show increases in power from baseline, with small additional random noise, then these two vectors should have high measured reinstatement. Pearson's correlation, a centered version of cosine similarity, would give a low correlation in this case because of the noise fluctuations, while the cosine similarity would be high, consistent with our interpretation of reinstatement. Thus, for each trial, n, we generate a temporal map of reinstatement values:

$$C_n(i,j) = \frac{\vec{E}_i \cdot \vec{R}_j}{\|\vec{E}_i\| \|\vec{R}_j\|} \tag{1}$$

where $C_n(i,j)$ corresponds to the reinstatement of neural activity across all electrodes and all frequencies between encoding epoch i and retrieval epoch j during trial n. For every patient, we compute the reinstatement maps separately for all correct, intrusion, and pass trials.

To calculate the reinstatement associated with every individual feature, k, we followed a similar procedure. In this case, we constructed a vector comprised of the z-scored power for that particular feature across all trials, $n = 1 \ldots N$. Thus, for every encoding temporal epoch, i, and every retrieval temporal epoch, j:

$$\vec{E}_i = [z_1(i) \ \ldots \ z_N(i)]$$

$$\vec{R}_j = [z_1(j) \ \ldots \ z_N(j)]$$

where $z_n(i)$ is the z-transformed power of feature k for trial $n = 1 \ldots N$ in temporal epoch i. For N trials, we thus create a vector containing N values for each feature for each time point. We then calculate reinstatement of that particular feature by calculating the cosine similarity between all encoding and retrieval vectors $\vec{E}_i$ and $\vec{R}_j$ for all pairs of encoding and retrieval temporal epochs. Thus, for every feature, k, we generate a reinstatement map, $C_k(i, j)$, corresponding to the reinstatement of that individual feature between encoding epoch i and retrieval epoch j across N trials. For every participant, we compute the reinstatement maps for each feature for all correct trials and for all incorrect trials.

6.2 Reinstatement Results

To assess differences in reinstatement between trial types across participants, we computed the average reinstatement for each encoding-retrieval time pair across trials for each participant and then computed the average of these reinstatement maps across participants (Fig. 6, see Fig. 7 for single participant examples). We only compared reinstatement between different trial types for an individual participant if both trial types included a minimum of ten trials each. We used a nonparametric clustering-based procedure [33] to assess whether there was a statistical difference in the reinstatement maps between trial types across participants. This procedure identifies contiguous time windows where reinstatement differs between trial types using a permutation procedure while avoiding *a priori* assumptions about particular temporal regions and correcting for multiple comparisons.

For every patient, we calculated the reinstatement map across trials for each trial type. For group analysis (Fig. 6), we computed the average reinstatement map for each individual, and then, for each encoding-retrieval time pair, we computed a p-

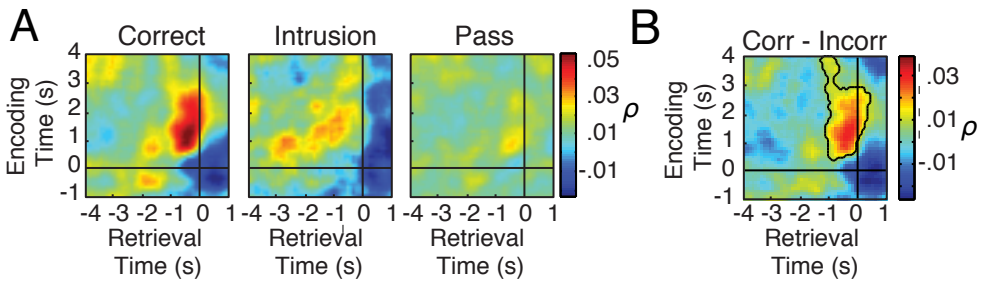

Figure 6 (*A*) Mean reinstatement across all participants during correct, intrusion, and pass trials. Black lines indicate the time of word presentation during encoding and the time of vocalization during retrieval. (*B*) The difference in mean reinstatement between correct and incorrect trials across all participants. We defined a temporal region of interest (tROI) for subsequent analyses as all encoding-retrieval time pairs that exhibited significant differences between the two trial types, outlined in black.

value and a t-statistic using a paired t-test comparing the distributions of averaged reinstatement values across patients between trial types. For individual participant analysis, we determined p-values for every encoding-retrieval time pair by using an unpaired t-test comparing the distributions of reinstatement values between trial types. In all analyses, we then permuted trial labels 1000 times, and for each permutation, calculated the reinstatement map for the surrogate trials from each trial type and repeated the t-test between trial types for each encoding-retrieval pair. For individual feature analysis, we assigned a p-value to every encoding-retrieval time pair by comparing the true difference in reinstatement between correct and incorrect trials to the distribution of differences found in the permuted cases for that time pair. For all analyses, in the true case and each of the permuted cases, we identified contiguous clusters of encoding-retrieval time pairs in which the p-value was below 0.05. For each cluster, we computed a cluster statistic by taking the sum of the t-statistic across all encoding-retrieval time pairs in that cluster. In this manner, large clusters statistics can arise either from large differences in reinstatement between trial types that extend over a short duration or from weaker differences that persist over longer time periods. In each permutation (n=1000), we saved the largest observed cluster statistic, thus generating an empiric distribution of maximum cluster statistics that

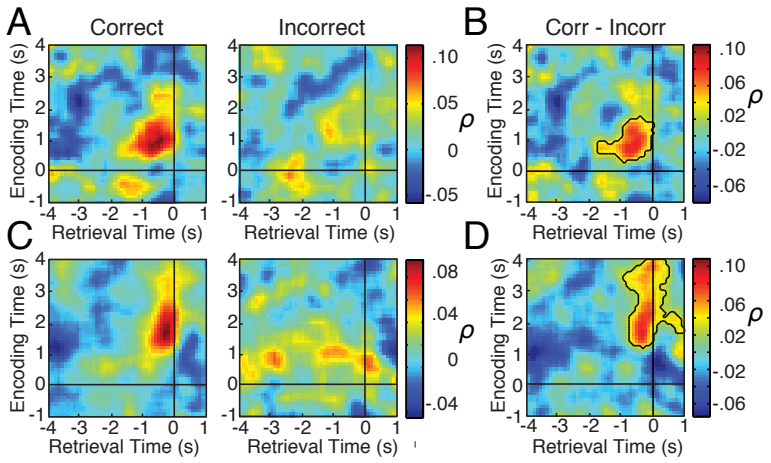

Figure 7 Reinstatement in individual participants. (*A*) Mean reinstatement for a single participant during all correct and incorrect trials. Color represents the mean reinstatement for each encoding-retrieval time pair, calculated by including all features for that participant across all trials. Black lines indicate the time of word presentation during encoding, and the time of vocalization during retrieval. (*B*) Difference in mean reinstatement between correct and incorrect trials for this participant. Encoding-retrieval time pairs exhibiting significant differences between the two conditions are outlined in black. (*C*) Mean reinstatement for a second participant during all correct and incorrect trials. (*D*) Difference in mean reinstatement between correct and incorrect trials for this participant.

would arise by chance. To generate a p-value for each cluster observed in the true dataset, we compared the position of the true cluster statistic to the distribution of maximum cluster statistics from the permuted cases. Clusters were determined to be significant if their p-value calculated in this manner was less than 0.05.

When we examined reinstatement during the retrieval period time locked to vocalization across participants, we found significantly greater reinstatement during correct trials compared to both pass trials and intrusion trials (Fig. 6; Fig. 8). We identified temporal encoding and retrieval epochs where reinstatement during correct trials was significantly greater than during incorrect trials, and defined this as our temporal region of interest (tROI) for further analyses ($P < 0.001$, permutation test; black outline, Fig. 6B). In a post-hoc analysis, we found that the mean reinstatement over the tROI, averaged across participants, was significantly greater during correct compared to both intrusion and pass trials ($t(29) = 5.19$, $P = 0.00001$ for correct vs pass, paired, two-tailed; $t(24) = 4.47$, $P = 0.0002$ for correct vs intrusion). We

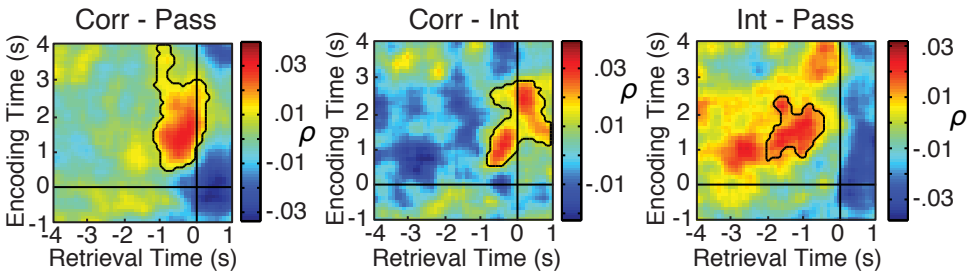

Figure 8 Differences in reinstatement between trial types. The average difference in mean reinstatement between correct and pass trials (*left*), between correct and intrusion trials (*middle*), and between intrusion and pass trials across all participants using all features, time locked to vocalization. Encoding-retrieval time pairs exhibiting significant differences between the two conditions are outlined in black.

identified the encoding-retrieval time pair within the tROI for each participant that exhibited the greatest reinstatement during correct trials and found that the peak encoding time occurred 1.68 ± 0.17 seconds after the presentation of the study pair, and the peak retrieval time occurred 0.44 ± 0.064 seconds prior to vocalization.

Reinstatement during memory encoding and retrieval may reflect trial-specific reinstatement of neural activity or general encoding and retrieval mechanisms. To investigate these possibilities, we compared the reinstatement observed during correct trials to reinstatement computed when we shuffled trial labels. We first shuffled the labels of all retrieval periods and calculated reinstatement between the true correct encoding periods and the shuffled retrieval periods. We found that the average reinstatement in the tROI originally observed with our data was significantly greater than reinstatement in this shuffled distribution ($t(29) = 8.50$, $P < 10^{-8}$, paired, two-tailed; Fig. 9A,B, Shuffle All). We next shuffled the labels of only the correct retrieval periods and calculated reinstatement between the true correct encoding periods and the shuffled correct retrieval periods. If reinstatement reflects a general encoding and retrieval process, then reinstatement calculated using shuffled correct retrieval periods should be identical to that observed using the original correct trials. Instead, we found that mean reinstatement in the tROI was significantly less when we used the shuffled compared to the true correct trials ($t(29) = 6.58$, $P < 10^{-6}$, paired,

two-tailed; Fig. 9A,B, Shuffle All Correct). When we restricted shuffling to only swap retrieval periods from adjacent correct trials, we also found that mean reinstatement in the tROI during correct trials was significantly greater than that calculated using the shuffled adjacent correct trials ($t(29) = 4.81$, $P = 0.00004$, paired, two-tailed; Fig. 9A,B, Shuffle Adjacent Correct).

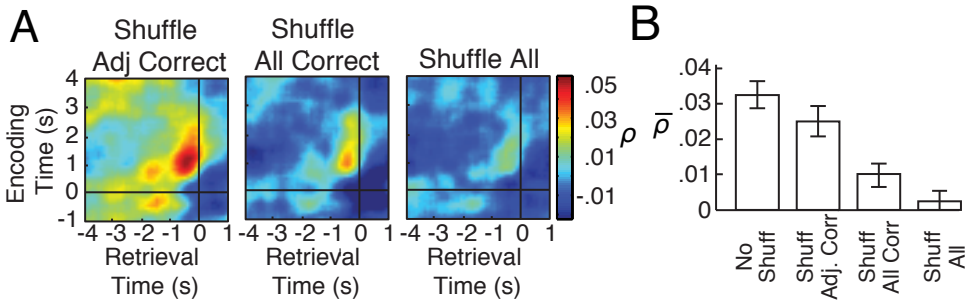

Figure 9 (*A*) Mean reinstatement during correct trials across all participants when shuffling neighboring correct retrieval periods (*left*), all correct retrieval periods (*middle*), or all retrieval periods from all trial types (*right*). (*B*) Mean reinstatement across all participants, averaged over the tROI for each participant, during each shuffled permutation shown in (A). Error bars represent SEM across participants.

We investigated the differential effect of shuffling on correct versus incorrect trials (Fig. 10). On average, there was a larger decrease in reinstatement over the tROI for shuffling correct trials (0.0227 ± 0.0034) versus shuffling incorrect trials (0.0189 ± 0.0035), but this difference was not statistically significant ($t(29) = 1.25$, $P = 0.23$, paired, two-tailed). As with the correct trials, the observed decreases with shuffled incorrect trials suggests that the extent of reinstatement observed during incorrect trials was largely related to the elapsed time between study and test.

The decreases in mean reinstatement observed with greater shuffling may be related to the amount of elapsed time between the true encoding and shuffled retrieval period. We therefore examined whether neural activity during an experimental session exhibited a slow temporal drift. We chose an identical temporal epoch from every encoding event (0 - 500 ms before stimulus presentation), and calculated the cosine similarity of neural activity between all pairs of the selected time epochs. We aver-

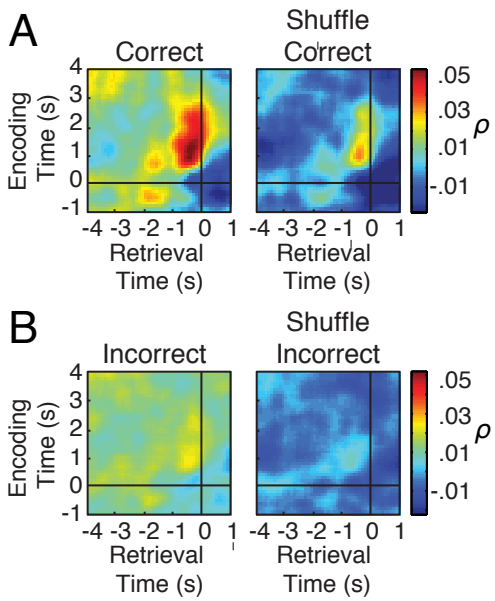

Figure 10 Differential effect of shuffling on correct vs. incorrect trials. (*A*) Mean reinstatement during correct trials across all participants (*left*) and when shuffling all correct retrieval periods (*right*). (*B*) Mean reinstatement during incorrect trials across all participants (*left*) and when shuffling all incorrect retrieval periods (*right*). For (A) and (B), black lines indicate the time of word presentation during encoding and the time of vocalization during retrieval.

aged the resulting values over all pairs separated by the same amount of time within and then across participants and binned the data separately into five and 60 second bins (Figure 11A,B). Because five second time bins approximately correspond to the timing of individual item presentation, the resulting data reflects the neural similarity between items separated by different lags within the same list. We examined the slopes of the resulting regressions and found that neural similarity was significantly related to the time between experimental events on both small and large time scales (five second bins, slope $= -.0141 \pm .0021$, $t(29) = -6.71$, $P < 10^{-6}$, one-sample, two-tailed; 60 second bins, slope $= -.0053 \pm .0011$, $t(29) = -4.67$, $P = 0.00006$).

We hypothesized that reinstatement during correct trials principally arose because neural activity jumped back in time to reinstate context present during encoding. To directly examine this, we identified the serial position during encoding of each correct trial and paired the associated retrieval period with all encoding periods from all serial

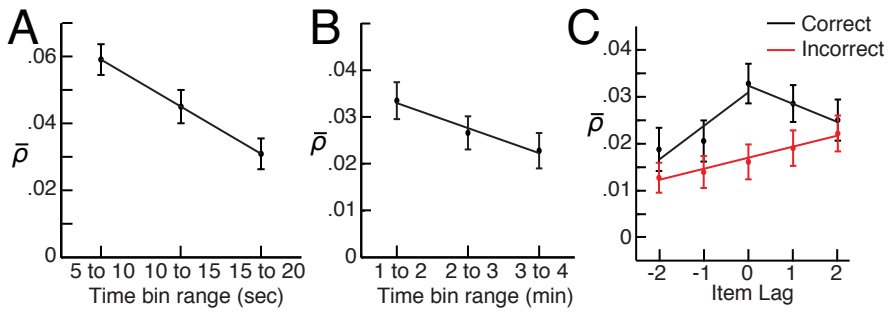

Figure 11 Neural reinstatement is shaped by temporal context. (*A*) Temporal autocorrelation of distributed oscillatory power during encoding. Plot shows the cosine similarity of neural activity between all pairs of time epochs from 0 - 500 ms before stimulus presentation during encoding. Data is shown for events separated by different lengths of time, separated into 5 second bins. (*B*) Same as (*A*), with data separated into 60 second bins. (*C*) Mean tROI reinstatement across all participants between each retrieval period and each encoding period from the associated study list (*black* correct, *red* incorrect). For (A-C), solid lines indicate the best fit regression line averaged across participants, and error bars represent SEM across participants.

positions within the same study list. In this manner, a true pairing would have a lag of 0, while a lag of 1 (-1) would correspond to a pairing between the true retrieval period and the encoding period of the subsequent (previous) item in the study list. We computed the average tROI reinstatement for each pairing across participants. If the reinstatement of temporal context governs the observed similarity, then neural activity during retrieval should be most similar to activity during the encoding period of the same pair, and should fall off with both positive and negative lag. Conversely, if temporal drift accounts for the observed similarity, then neural activity during retrieval should be most similar to activity during the most recent encoding period (i.e., most positive lag).

For correct trials, we found that each retrieval period exhibited maximum reinstatement with the true encoding period, and that when paired with other encoding periods from the same list, mean reinstatement decreased with both positive and negative lag (Fig. 11C). We regressed mean reinstatement with lag and found an average slope that was positive across participants when retrieval periods were paired with previous encoding periods (negative lag, slope = $.0068 \pm .0018$), and an av-

erage slope that was negative when retrieval periods were paired with subsequent encoding periods (positive lag, slope $= -.0037 \pm .0014$). The decrease in mean reinstatement observed with positive and negative lags during correct retrievals suggests that successful reinstatement was shaped by the reinstatement of temporal context during encoding. We performed the same analysis using incorrect trials, and found that mean reinstatement across participants exhibited a progressive increase with encoding lag (slope $= .0023 \pm .0006$). The increase with encoding lag suggests that measured reinstatement during incorrect trials is governed most by temporal proximity to the item at test. To directly assess whether the reinstatement of temporal context during encoding was particular to correct retrieval, we compared the relation between positive lag (0 - 2) and mean reinstatement during correct trials to the relation between positive lag and mean reinstatement during incorrect trials. We found that the relation between reinstatement and positive lag was significantly different between correct and incorrect trials ($t(29) = -3.41$, $P = 0.0019$, paired, two-tailed).

To assess the contributions of different frequency bands to reinstatement, we calculated the mean reinstatement during correct trials over our tROI using only features from each frequency band and found a significant effect of frequency on mean reinstatement ($F(4, 145) = 3.34$, $P = 0.012$, one-way ANOVA; Fig. 12A). Pair-wise t-tests showed that reinstatement was significantly greater in both high gamma and theta as compared to alpha and beta ($t(29) > 2.34$, $P < 0.03$, paired, two-tailed), but there was no difference between theta and high gamma ($t(29) = -1.26$, $P = 0.22$). We next examined the temporal dynamics underlying the contributions of different frequencies to reinstatement during correct trials. First, we calculated the mean reinstatement for all encoding-retrieval time pairs across participants using only theta or high gamma features and observed that the time periods exhibiting significant differences in reinstatement between correct and incorrect trials varied between frequency bands (Fig. 12B). We then averaged mean reinstatement, calculated using

only features drawn separately from each frequency band, over all encoding temporal epochs defined in the tROI. High gamma and theta features exhibited the largest rises in reinstatement during correct trials, peaking immediately before vocalization (Fig. 12C *top*). We then averaged mean reinstatement over all retrieval epochs defined in the tROI separately for each frequency band. First high gamma, and then theta, exhibited peaks in reinstatement around one second following word presentation during encoding (Fig. 12C *bottom*).

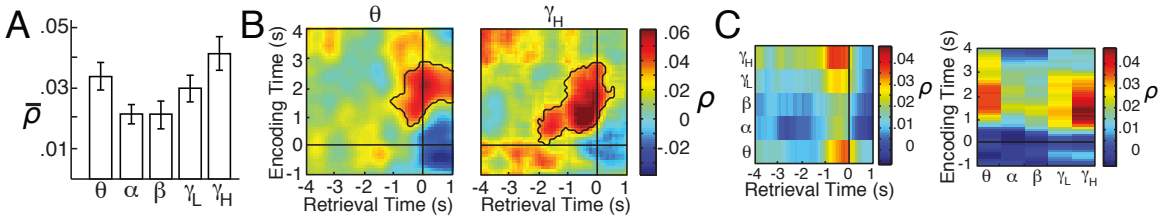

Figure 12 Neural reinstatement is mediated by temporally precise theta and high gamma frequency activity in specific anatomic locations. (*A*) Mean reinstatement for all features in each frequency band averaged over the tROI for each participant. Error bars represent SEM across participants. (*B*) Mean reinstatement of correct trials across all participants for theta (*left*) and high gamma (*right*) frequency bands. Regions of significant difference between correct and incorrect trials are outlined in black. Black lines indicate time of vocalization during retrieval and time of pair presentation during encoding. (*C*) Mean reinstatement for each frequency, averaged over all encoding (*top*) or retrieval (*bottom*) time periods defined by the tROI, for all retrieval (*top*) or encoding (*bottom*) times. Color represents the average mean reinstatement for each frequency across participants. The black lines indicate the point of vocalization during retrieval (*top*) and the point of stimulus presentation during encoding (*bottom*).

As reinstatement seemed to be principally mediated by theta and high gamma frequencies, we visualized the anatomic regions, separately for each frequency band, that contained features exhibiting significant differences in mean reinstatement over the tROI between correct and incorrect trials across participants. Only regions that had electrodes from five or more participants were included in this analysis (Fig. 13). We rendered all spatial regions exhibiting significant differences across participants ($P < 0.004$, permutation procedure, FDR corrected), where color intensity represents mean reinstatement during correct trials for that spatial region and frequency (Fig. 13). We found significantly greater reinstatement for correct versus incorrect trials

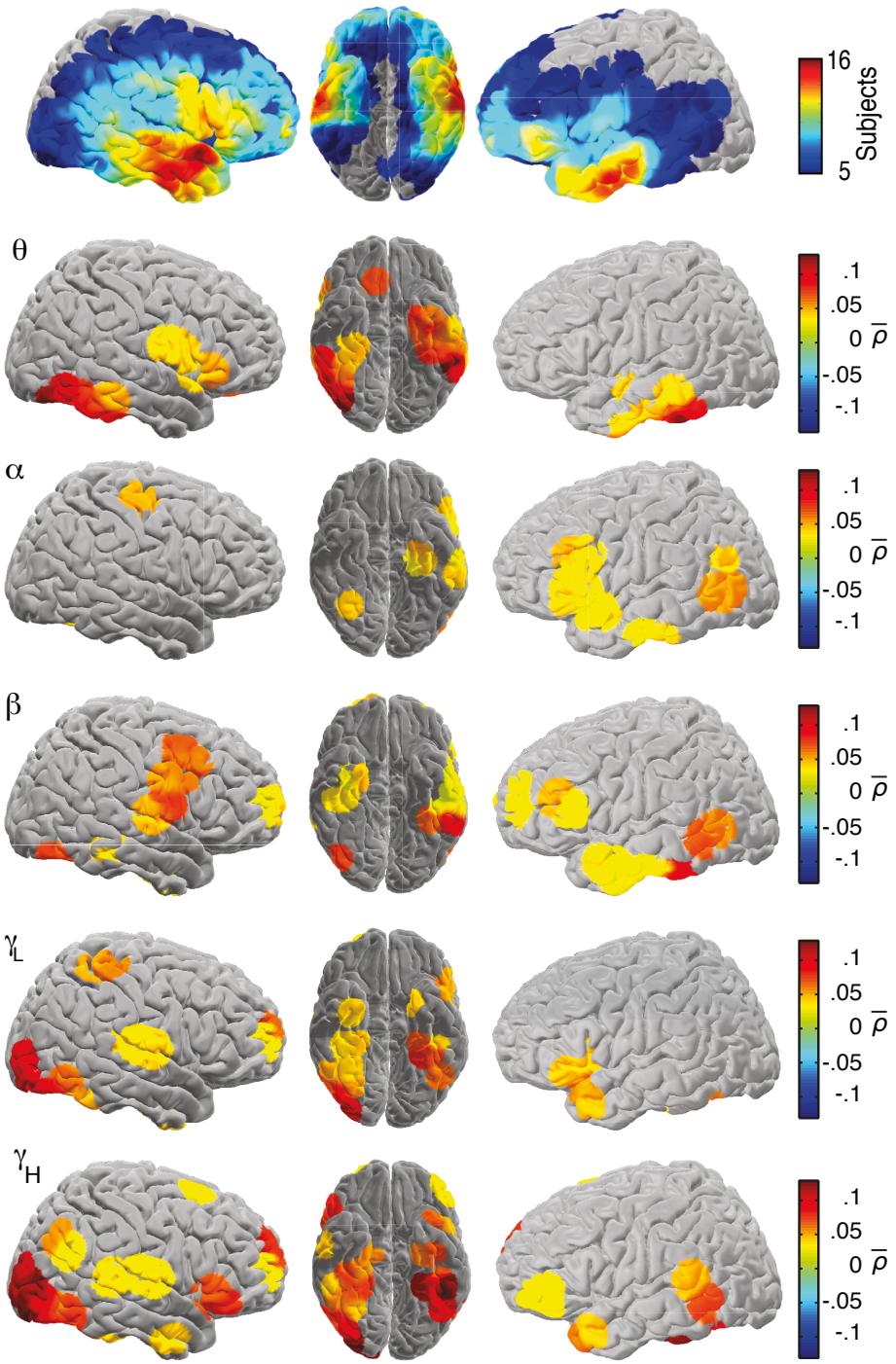

Figure 13 Neural reinstatement localizes to specific anatomic locations. (*top*) Electrode coverage across all 32 participants. Bottom 5 rows show anatomic regions exhibiting significant differences in reinstatement for all theta, alpha, beta, low gamma and high gamma features betweencorrect and incorrect trials across all participants. Only spatial ROIs with features exhibiting significant differences are colored. Color represents the mean reinstatement across correct trials, averaged across all participants, for all features contributing to each ROI.

in the inferior temporal lobes bilaterally in both theta and high gamma frequency bands. We also found significant reinstatement in the right ventrolateral prefrontal cortex in theta, and bilaterally in high gamma. In high gamma, we also found areas of significant reinstatement that included the prefrontal cortex, the right supplementary motor area, the right temporo-parietal junction, the left posterior lateral temporal lobe, and the right occipital lobe.

To determine the spatial distribution of reinstatement, we first created 888 spatial regions of interest (sROI) evenly spaced every 9.64 ± 0.02 mm on the cortical surface of a Montreal Neurological Institute N27 standard brain. We also created an additional 741 sROI's within the brain parenchyma, evenly spaced every 9.58 ± 0.02 mm. We grouped spatially similar electrodes from different participants by identifying electrodes that fell within 12.5 mm of each sROI.

In order to determine whether a particular sROI exhibited significant differences in reinstatement between correct and incorrect trials across participants, we used a permutation procedure. We calculated the difference in mean reinstatement over the tROI between correct and incorrect trials for each feature for each participant. We grouped similar spatial features within each sROI by taking the average of those differences within participants, and then for each sROI calculated the average difference across all participants with features in that sROI. We then permuted the trial labels for the trial types 1000 times and calculated a distribution of average differences across participants for each sROI. To generate a p-value for the differences in reinstatement for a given sROI, we determined the position of the true average difference in the distribution of permuted average differences. We used a false discovery rate (FDR) procedure with $q = .05$ applied to all p-values from all sROI's to correct for multiple comparisons [16]. To determine whether an sROI exhibited significant differences in reinstatement in an individual, we did not use the across-participant tROI as each individual exhibits unique patterns of reinstatement. Instead, we used the

nonparametric clustering procedure to identify a significant temporal cluster for each sROI, with the constraint that the peak value of reinstatement within that cluster must occur within 4 sec of word presentation during encoding and between 3 sec prior to 1 sec after vocalization during retrieval.

To visualize the spatially precise changes in reinstatement across all participants, we identified sROI's that exhibited a significant ($p < 0.004$, FDR corrected across all sROI's) difference in reinstatement between correct and incorrect trials for all features in each frequency band. Cortical topographic plots were rendered by assigning each vertex in the 3D rendered image of the standard brain a weighted average of the mean correct reinstatement of each sROI that includes that vertex. Weighted reinstatement values for each vertex within a single sROI were assigned by convolving the value of the sROI with a three dimensional Gaussian kernel (radius = 12.5 mm; $\sigma = 4.17$ mm) with center weight 1. The value of each vertex on the standard brain was therefore computed by taking the average of the weighted contributions of all overlapping sROI's. We projected these vertex values onto the standard brain using information from the WFU PickAtlas toolbox [31] (Fig. 13). All colored regions, independent of color intensity, identically indicate one-tailed significance at the $p < 0.004$ level as determined by the permutation procedure. For individual topographic plots, all colored regions indicate significance at the $p < .05$ level as determined by the permutation procedure. Intensity varied as a function of reinstatement strength and with the standard deviation of the Gaussian kernel, which was used purely as a visualization technique to identify areas in which many contiguous sROI's exhibited statistically significant changes in reinstatement.

Because we observed that the specific time periods of reinstatement varied between participants, between frequency bands, and between lobes, we hypothesized that the time course of reinstatement for individual neural features would also vary. We examined differences in reinstatement between correct and incorrect trials for in-

dividual features in individual participants. For each participant, we visualized the anatomic regions and frequency bands that contained features exhibiting significant differences between correct and incorrect trials (Fig. 14A). On average, 11.3% ± 0.9% of features demonstrated significant differences in reinstatement for a given participant ($P < 0.05$, permutation test). The encoding-retrieval time regions exhibiting significant differences in reinstatement varied among features (Fig. 14A). For each feature, we identified the encoding-retrieval time pair that exhibited the maximum reinstatement during correct trials within these identified time regions. When we examined the timing distribution of all significant theta and high gamma features, we found that peak encoding times were significantly greater for theta features than for high gamma features ($t(414) = 2.90$, $P = 0.004$, unpaired, two-tailed; Fig. 14B). Peak retrieval times did not significantly differ between theta and high gamma features ($t(414) = 0.37$, $P = 0.71$).

To further investigate the differences in the timing of reinstatement, for each participant, we calculated the difference between the peak encoding and peak retrieval times for all significant theta and high gamma features. We first only included individual electrodes that demonstrated *both* significant theta and high gamma reinstatement, and found that peak high gamma activity significantly preceded peak theta activity during encoding across participants (Fig. 14C; $t(20) = 3.46$, $P = 0.003$, one-sample, two-tailed). Notably, during retrieval, we found no difference in peak theta and high gamma times ($t(20) = 1.49$, $P = 0.15$). The difference we observed in encoding times between high gamma and theta was significantly greater than the difference in retrieval times in these electrodes ($P = 0.039$, permutation test), which were mostly located in the left posterior temporal lobe, the right fronto-temporal region, and the bilateral temporo-occipital areas (Fig. 14D). We next included all electrodes that exhibited *either* significant high gamma or theta reinstatement, and in each participant computed the average peak encoding and retrieval times for each

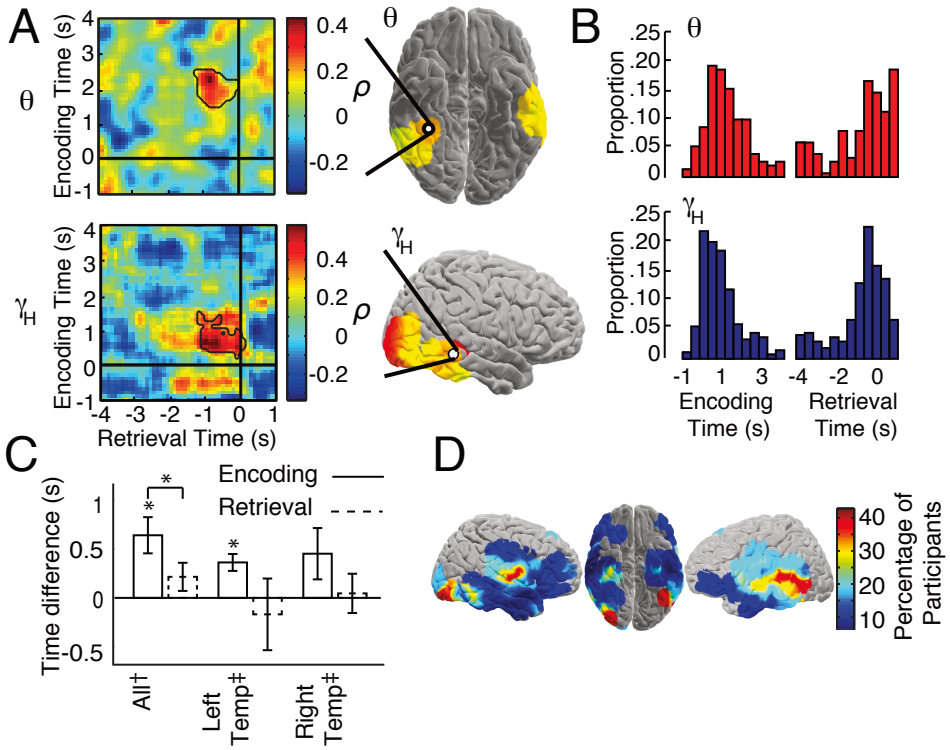

Figure 14 High gamma reinstatement precedes theta during encoding but not during retrieval. (*A*) Mean reinstatement during correct trials of one theta (*top*) and one high gamma (*bottom*) feature in a single participant. Encoding-retrieval time pairs exhibiting significant differences between correct and incorrect trials are outlined in black. The anatomic location of the individual features are represented as white circles on the topographic plots. Topographic plots represent anatomic regions exhibiting significant differences in reinstatement between correct and incorrect trials for all theta and high gamma features for this participant. (*B*) Distributions of peak encoding and retrieval times across all participants for theta (*top*) and high gamma (*bottom*) features exhibiting significant differences in reinstatement between correct and incorrect trials. (*C*) Differences in peak reinstatement times between theta and high gamma features during encoding and retrieval. Differences in peak time are shown for all electrodes that exhibit significant reinstatement in *both* theta and high gamma on the *left* (single dagger). Peak time differences are shown for all electrodes in the left and right temporal lobes that exhibit *either* theta or high gamma reinstatement are shown in the *middle* and on the *right*, respectively (double dagger). Asterisks indicate a difference between peak theta and high gamma time that is significantly greater than zero ($P < 0.05$) or a significant difference between encoding time differences and retrieval time differences ($P < 0.05$). Error bars represent SEM across participants. (*D*) Percent of participants with electrodes in a given spatial ROI that exhibited significant reinstatement in *both* theta and high gamma.

frequency for different anatomic regions. Across participants, the left temporal lobe exhibited a significant difference in peak time of high gamma and theta reinstatement during encoding ($t(10) = 4.18$, $P = 0.002$, one-sample, two-tailed; Fig. 14C; right medial temporal lobe $t(5) = 2.74$, $P = 0.041$, not shown; other regions $P > 0.05$). Conversely, during retrieval, we found no anatomic region that exhibited significant differences in the peak time of theta and high gamma reinstatement ($P > 0.05$).

We examined reinstatement during the retrieval period time locked to cue presentation and found that reinstatement during correct trials was significantly greater than during incorrect trials, and the time where these differences were significant extended between 1 to 2 sec during the encoding period and between 0.5 and 1.3 sec after word presentation ($P = 0.013$, permutation test; Fig. 15A,B). However, we also found that reinstatement of neural activity was present in all three trial types when locked to cue presentation during retrieval (Fig. 15A). We found time regions with significant differences in reinstatement between correct and pass trials and between intrusion and pass trials, although we did not find significant differences between correct and intrusion trials (Fig. 15C). We computed the mean reinstatement over the time region that exhibited significant differences between correct and incorrect trials for each participant and averaged across participants for each trial type. We found that the mean reinstatement over this time region was significantly greater during correct compared to pass trials ($t(29) = 6.53$, $P < 10^{-6}$ for correct vs pass, paired, two-tailed; $t(24) = 0.25$, $P = 0.40$ for correct vs intrusion; Fig. 15D).

To investigate why the differences in reinstatement were greater when time locked to vocalization as compared to when time locked to cue presentation, when reinstatement appeared in all three trial types, we examined the z-scored power during every temporal encoding and retrieval epoch for every electrode and for every frequency. While we found a number of electrodes that exhibited changes in activity during successful encoding and retrieval (Fig. 16A), we also observed a number of

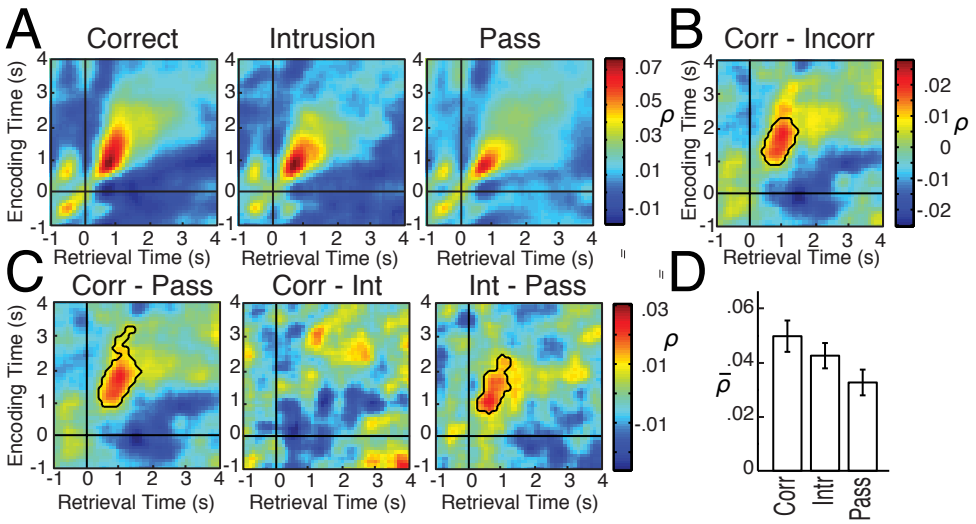

Figure 15 Neural reinstatement time-locked to cue presentation. (*A*) Mean reinstatement across all participants during all correct, intrusion, and pass trials, time locked to cue presentation during retrieval. Color represents the average mean reinstatement across participants for each encoding-retrieval time pair, calculated by including all features for each participant. Black lines indicate the time of word presentation during encoding and the time of cue presentation during retrieval. (*B*) The average difference in mean reinstatement between correct and incorrect trials across all participants, time locked to cue presentation. Time pairs with significant differences between the two conditions are outlined in black. (*C*) The average difference in mean reinstatement between correct and pass trials (*left*), between correct and intrusion trials (*middle*), and between intrusion and pass trials (*right*) across all participants using all features, time locked to cue presentation. Encoding-retrieval time pairs exhibiting significant differences between the two conditions are outlined in black. (*D*) The mean reinstatement over the time region identified in (B) for each participant, averaged across participants for correct, intrusion, and pass trials. Error bars represent SEM.

electrodes that demonstrated increases in oscillatory power time-locked both to word presentation during encoding and to cue presentation during retrieval (Fig. 16B). If reinstatement were solely due to reactivation of these visually responsive features, then we would expect to see no difference in reinstatement between correct and incorrect trials as these features would equally respond to cue presentation in both conditions. In this scenario, the observed differences in reinstatement when locked to vocalization could be explained by differences in response time distributions (Fig. 2B) since correct trials, having faster and more consistent response times, would exhibit cue-evoked visual responses that are more aligned immediately prior to vocalization. We note, however, that even when locked to cue presentation, there still remain tem-

poral regions with significantly greater reinstatement during correct trials compared to pass trials, suggesting that these regions reflect reinstatement related to memory encoding (Fig. 15C).

To account for differences in response time distributions between correct and incorrect trials, we computed the mean reinstatement after assigning response times of incorrect trials with response times that were drawn from the distribution of correct response times (Fig. 16C). This eliminates any discrepancies in reinstatement that may arise from differences in response time distribution, and aligns any visually evoked reinstatement in both conditions. We found significantly greater reinstatement during correct trials compared to incorrect trials extending from 1 sec before to 0.2 sec after vocalization and from 0.8 to 2.8 sec following word presentation during encoding ($P = 0.008$, permutation test; Fig. 16D).

In order to identify features that respond to a visual stimulus, we examined reinstatement during pass trials time-locked to cue presentation (Fig. 15A). Because participants did not successfully recall items during these trials, this reinstatement represents the reinstatement of visually evoked activity. We identified which temporal region exhibited visually evoked reinstatement by comparing the distribution of reinstatement across participants at each encoding-retrieval time pair during pass trials to the average reinstatement across all encoding-retrieval time pairs during pass trials using a t-test. We identified any encoding-retrieval time pair with $P < 9.81e-6$ (Bonferroni corrected across all possible encoding-retrieval time pairs) as exhibiting a significant visual response. The maximum cue-locked retrieval time that exhibited significant reinstatement during pass trials was 1300 ms after cue presentation. We defined this time as the extent of visually evoked reinstatement.

We then identified individual features that exhibited a visual response by examining changes in z-scored oscillatory power between cue presentation and 1300 ms during retrieval. Any feature that exhibited a significant ($P < 0.05$, unpaired, two-

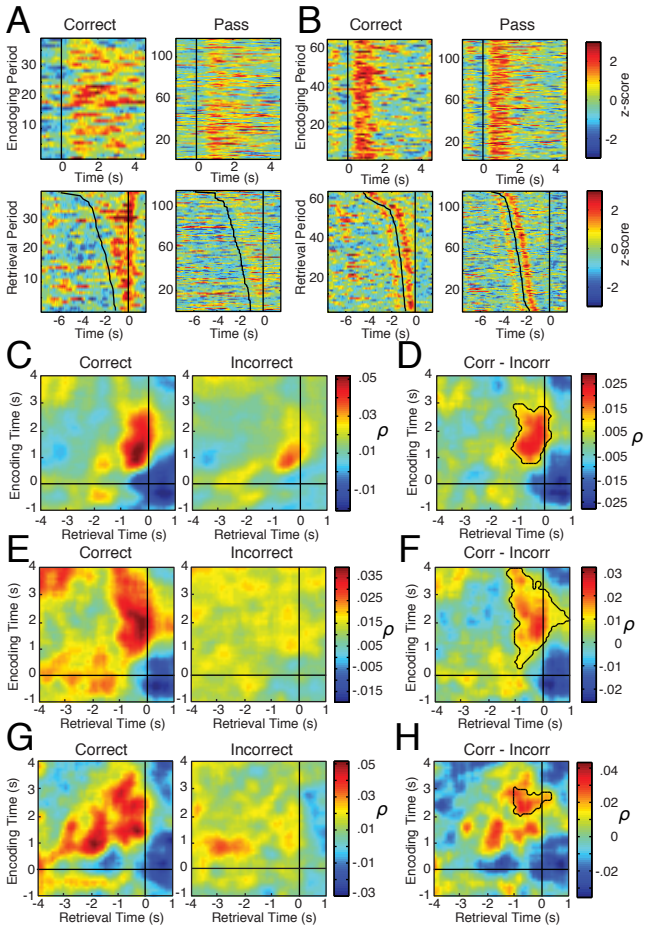

Figure 16 Significant differences in reinstatement persist when accounting for visual response confounds. (*A*) *z*-scored high gamma power for a single electrode in the left temporal lobe during correct (*left*) and pass (*right*) trials during encoding (*top*) and retrieval (*bottom*). Note that activity is not time locked to visual presentation in correct retrieval trials. (*B*) Same as (A) for an electrode in the left occipital lobe. Note that activity is locked to visual presentation in all trials. For (A) and (B), black line during encoding represents the time of word presentation, black lines during retrieval represent the time of cue presentation and subsequent vocalization for each trial, and trials are sorted by response time during retrieval. (*C*) Mean reinstatement calculated after assigning all incorrect trials a response time randomly drawn from the distribution of response times observed during correct trials. (*D*) The difference in mean reinstatement between correct and incorrect trials from (C). (*E*) Mean reinstatement calculated after removing features that exhibited a significant visual response during pass retrieval trials. (*F*) The difference in mean reinstatement between correct and incorrect trials from (E). (*G*) Mean reinstatement calculated after removing all trials with a response time less than 2.3 seconds. (*H*) The difference in mean reinstatement between correct and incorrect trials from (G). For (C)-(G), black lines indicate the time of word presentation during encoding, and the time of vocalization during retrieval. For (D), (F), and (H), encoding-retrieval time pairs exhibiting significant differences between the two conditions are outlined in black.

tailed t-test) change in mean z-scored oscillatory power from baseline over this time window (0 to 1300 ms during retrieval) was labelled as a visually responsive feature. We identified 177 ± 17 such features per participant. We removed these features and compared reinstatement time locked to vocalization between correct and incorrect trials using the remaining features (Fig. 16E). We found significantly greater oscillatory reinstatement during correct trials compared to incorrect trials extending from 1 sec before to 1 sec after vocalization and from 0.3 to 4 sec following word presentation during encoding ($P = 0.002$, permutation test; Fig. 16F).

To further isolate reinstatement unrelated to visual responses, we also discarded all trials with response times less than 2300 ms (allowing at least one additional second after any visually responsive effects during retrieval). This eliminates any advantage correct trials may have due to faster and more consistent response times as the remaining trials are associated with a significant time delay between visual presentation and vocalization. We compared reinstatement time locked to vocalization between the remaining 23 ± 8 correct trials and the 55 ± 7 incorrect trials (Fig. 16G). Mean reinstatement was again significantly greater during correct trials compared to incorrect trials from 1 sec before to 0.3 sec after vocalization and during 2 to 3 sec following word presentation during encoding ($P = 0.012$, permutation test; Fig. 16H).

To investigate the changes in spectral power that contribute to reinstatement in selected features that exhibited significant differences in reinstatement between correct and incorrect trials, we analyzed the average z-scored power during the encoding and retrieval epochs defined by the tROI. A positive (negative) average z-score indicates an increase (decrease) in power during successful encoding or retrieval. Power may either increase or decrease during both encoding and retrieval. For theta (high gamma) features that exhibit significant differences in reinstatement between correct and incorrect trials, $26.3 \pm 5.4\%$ ($50.3 \pm 5.2\%$) showed increases while $47.8 \pm$

6.3% (24.5 ± 4.2%) showed decreases during both encoding and retrieval. 25.5 ± 4.5% (25.2 ± 3.5%) of these identified features showed inconsistent changes between encoding and retrieval. Overall, theta features contributing to reinstatement most likely showed power decreases while high gamma features most likely showed power increases during both encoding and retrieval.

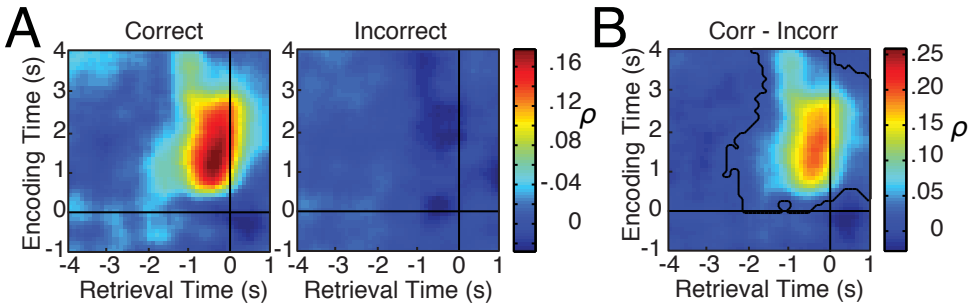

Figure 17 Reinstatement occurs with high fidelity using selected features. (*A*) Mean reinstatement of correct and incorrect trials across all participants including all features that exhibited a significant difference in reinstatement between correct and incorrect trials over the tROI. Black lines indicate time of vocalization during retrieval and time of pair presentation during encoding. (*B*) Difference in mean reinstatement between correct and incorrect trials including all significant features. Regions of significant difference are outlined in black.

Our data suggest that the reinstatement of neural activity observed in Figure 6 is mostly driven by a small set of individual features in specific spatial locations and in specific frequency bands. We identified these features by selecting all features that exhibited a significant difference between reinstatement across correct trials and reinstatement across incorrect trials over the tROI ($P < 0.05$, permutation test). We examined reinstatement of neural activity using only those identified features (Fig. 17A) and found, again, that there was significantly greater reinstatement during correct trials compared to incorrect trials ($P < 0.001$, permutation test; Fig. 17B). We identified the encoding and retrieval time that exhibited peak reinstatement using these selected features during correct trials, and found that the extent of reinstatement at this time was significantly greater than the reinstatement at the peak encoding and retrieval time for correct trials when using all features ($t(29) = 12.52$, $P < 10^{-12}$).

7 Timescale of Encoding and Retrieval Processes

Many of the analyses performed thus far have suggested that the progression of activity that occurs during retrieval occurs on a faster timescale than the progression of activity during encoding. Specifically, the evidence presented in Figure 10 suggests replay of activity occurs on a faster timescale during retrieval. The reinstatement map showing both encoding and retrieval time-locked to cue presentation (Figure 18) provides further evidence of this. In this plot, a black line is drawn along the y=x line to show that the reinstatement does not fall along this line as it would if the replay of activity during retrieval occurred on the same timescale as encoding. Instead, the reinstatement is elongated along the encoding axis. This suggests that the activity during retrieval occurs on a faster timescale than the activity during encoding.

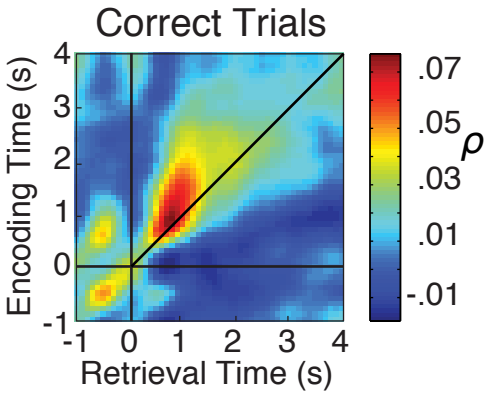

Figure 18 Reinstatement map for correct trials, where time 0 on the encoding axis is when the pair of words is presented on the screen. During retrieval, time 0 here is when the cue is presented on the screen. A black line is drawn along the y=x line to highlight the asymmetry of the reinstatement effect.

7.1 Time Warping Algorithm

In order to explicitly measure the differences in the timescales between encoding and retrieval, a time warping algorithm was applied to the average power trace for each brain region. This algorithm was applied to every spatial ROI and done independently for the high gamma and theta power traces. For each subject, for each electrode,

for each frequency band, the average power trace was computed for encoding and retrieval across all correct trials. Then, the average power traces were averaged across electrodes within each spatial ROI. This was done so that for each spatial ROI and frequency band, we have an average encoding trace from -0.5 sec to 4 seconds (time 0 is when the pair was presented) and an average retrieval trace from -0.5 sec to 4 seconds (time 0 is when the cue was presented). Now, the retrieval trace is stretched or compressed to best match the encoding trace from 0 to 2 sec. This is done by changing the timescale of the retrieval trace and then interpolating the retrieval trace so that it is sampled at the same points as the encoding trace. The optimal time scaling factor is determined as the time scaling factor that yields the maximum correlation between the encoding trace from 0 to 2 seconds and the warped retrieval trace. A time scale factor less than 1 indicates that the retrieval trace has been stretched (which would correspond to retrieval occurring on a faster timescale), and a factor greater than 1 indicates the opposite. The range of timescales that are searched is from 0.5 to 2. A timescale factor of 0.5 would mean that retrieval processes occurs twice as fast as encoding, while a factor of 2 would mean that retrieval occurs two times slower. An example of this algorithm being applied is shown in Figure 19.

The warping algorithm was applied to every region of the brain for theta power and high gamma power. However, this warping algorithm could be applied to any two arbitrary signals and would yield some optimal scale factor. In order to determine that a particular region exhibited patterns of spectral power that were similar enough between encoding and retrieval that one could be considered a warped version of the other, we performed a permutation test to establish that the maximum correlation achieved was significantly greater than what you would expect to see by chance. For each of 500 permutations, each region for each subject was randomly paired with another region from that subject. A region was considered to have achieved a significant correlation if the true correlation was greater than 95% of the permuted

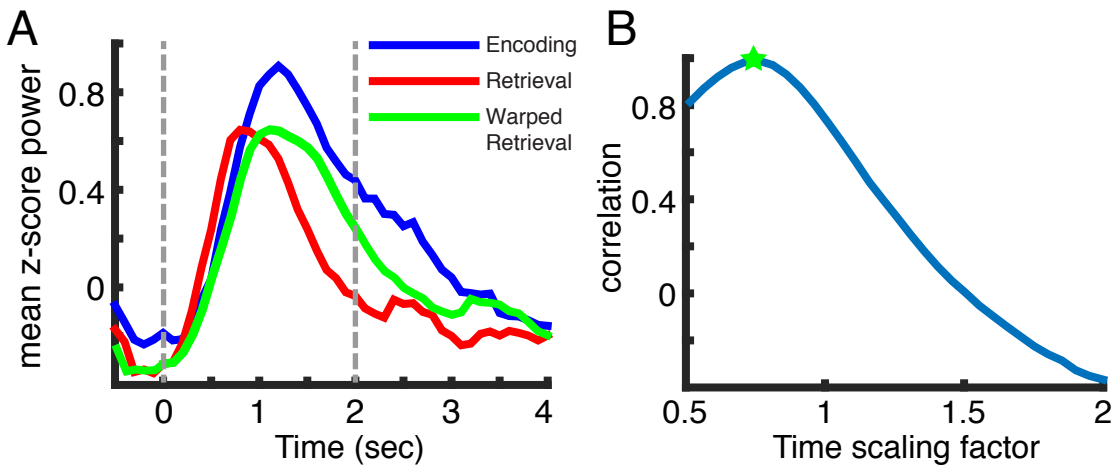

Figure 19 (A) For each brain region, a warping algorithm was applied to determine the optimal scaling factor to maximize the correlation of the power between encoding and retrieval. The retrieval trace is warped to best match 0-2 seconds during encoding. Dashed grey lines indicate this region. This is an example of high gamma power in the inferior temporal gyrus. (B) Here, the scaling factor that maximizes the correlation is 0.75 (indicated by green star), which means that retrieval occurs on a faster timescale than encoding. A scale factor greater than 1 would indicate that retrieval occurs on a slower timescale.

correlations. The average maximum correlations that were achieved for each region are shown in Figure 20A.

7.2 Timescale Analysis

Many brain regions were able to achieve significant correlations with the time warping function applied. For each subject, the average scale factor was computed for all the regions that achieved a significant correlation. Then the distribution of average scale factors was considered across subjects for both theta and high gamma power. For theta and high gamma, the mean scale factors across subjects were 0.92 ± 0.03 and 0.95 ± 0.02, respectively. The mean scale factors for theta and high gamma were significantly different from 1 (Theta: $P = 0.005$, t-test. High gamma: $P = 0.049$, t-test). Figure 20B (top) shows the scale factors for the regions that achieved a significant correlation across subjects. Figure 20B (bottom) shows the scale factors for regions that achieved a significant correlation across subjects and scale factors

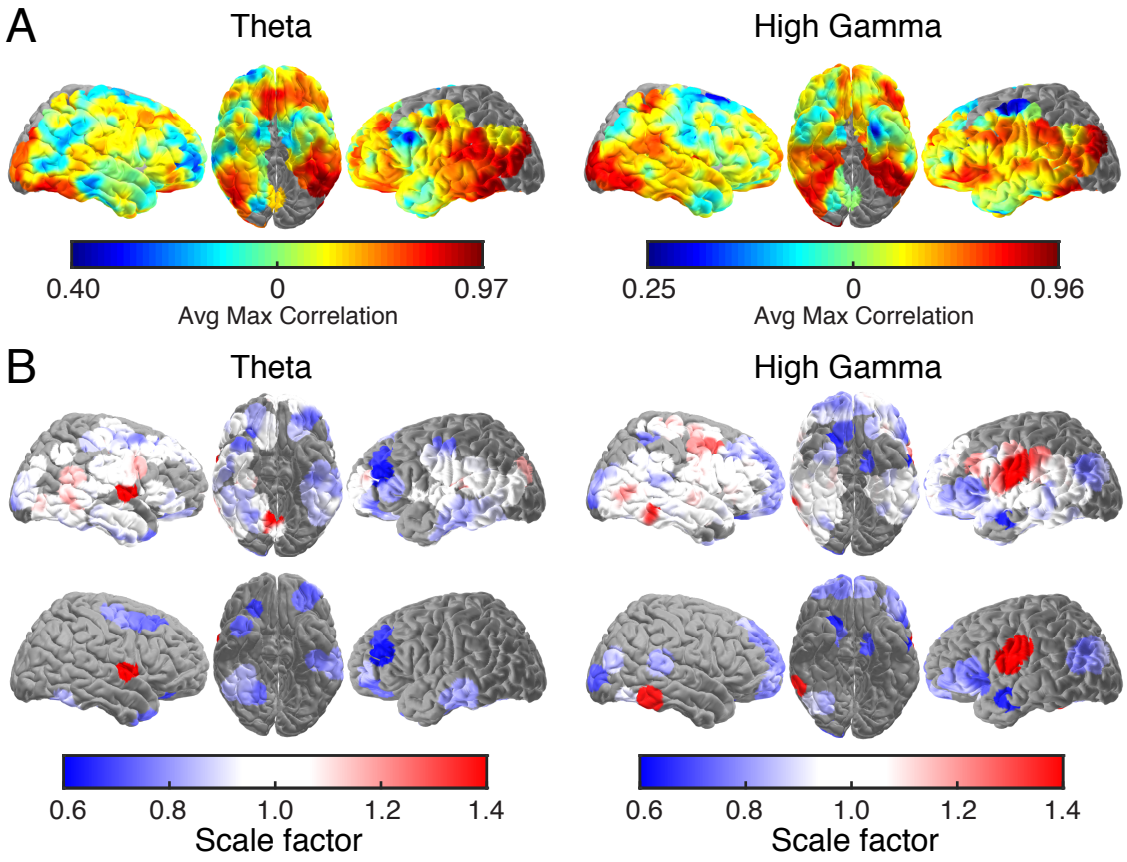

Figure 20 (A) The average correlation that was achieved with the optimal warping shown for each brain region for theta and high gamma. (B) The top row of brains shows the average scaling factors for regions that were able to achieve a significant correlation ($P < 0.05$, permutation test, uncorrected) with the warping function. The bottom row of brains shows the average scaling factors for regions whose scaling factors were significantly different from 1 ($P < 0.05$, t-test, uncorrected). Brains are shown for theta and high gamma.

that were significantly different from 1 ($P < 0.05$, uncorrected).

To ensure that the algorithm was not biased towards scale factors less than 1, we repeated the analysis reversing the roles of encoding and retrieval. Now, the average encoding trace was warped to best match the retrieval trace from 0 to 2 sec. This yielded a result consistent with the interpretation of the previous result. In this case, the mean scale factors across subjects for theta and high gamma were 1.07 ± 0.03 and 1.09 ± 0.03, respectively. The mean scale factors were significantly different from 1 (Theta: $P = 0.016$, t-test. High gamma: $P = 0.002$, t-test).

Our results demonstrate that on average, spectral dynamics are replayed on a

faster timescale during retrieval. We found that the speedup occurred for both theta and high gamma frequency bands. It is particularly interesting to note where the significant speedups are occurring in both frequency bands. For both theta and high gamma, we see a significant speedup in the left prefrontal cortex. For theta, we see speedups in the inferior temporal gyrus, and for high gamma, we see speedups in the frontal lobes. It is also worth noting that there are many brain regions that show replay of encoding activity during retrieval on the same timescale. Furthermore, there are a few regions that show a replay of activity on a slower timescale during encoding.

It may be the case that the regions that show a speedup of activity during retrieval are memory specific regions. During both encoding and retrieval, words are presented on the screen and must be read. The visual processing of words should occur on roughly the same timescale during encoding and retrieval. During retrieval, however, there is a single cue word that needs to be read, while during encoding, there are two words that need to be read. This could potentially account for some of the speed differences along the ventral processing pathway. Additionally, during encoding, the memory relevant areas in the left temporal and prefrontal cortices are activated during the later encoding period, presumably after the pair of words has been read and is being stored into memory. During retrieval, the cue causes the reactivation of these regions, which may cause them to be activated more quickly during retrieval compared to encoding. One additional observation is that we see a slower timescale of activity in some speech motor areas during retrieval. This may occur because subjects may be saying the pair of words to themselves as the words are shown on the screen during encoding. This would occur early in the encoding period, activating speech motor areas. During the later retrieval period, when the response is given, there is strong activation of the speech motor areas. This pattern of activity may explain the regions that show slower timescales of activation during retrieval.

Many previous studies have investigated the replay of neural activity in rodents as

it pertains to memory formation and retrieval [9,11,14]. These studies have primarily focused on hippocampal microelectrode recordings and have shown many examples of replay of neural spiking activity occurring on a faster timescale. It is not clear whether there is an underlying common mechanism between hippocampal spiking replay and the cortical spectral power replay we observe in this study. Further investigation is required to fully answer the question of why different brain regions show different time scaling factors. Additionally, future investigation is required to see if there are any behavioral correlates to the timescale on which retrieval occurs.

8 Reinstatement of Mutual Information Networks

We have explored the reinstatement of spectral power, and now we will investigate the functional connections that underlie these distributed patterns of spectral power, and whether the patterns of communication themselves are reinstated between memory encoding and retrieval. If distributed patterns of neural activity are reactivated during successful recall, and functional connections serve to bind neural activity across regions, then we hypothesized that such communication underlies the formation and reinstatement of these association networks. Among the several challenges in the study of human brain networks, however, are both how functional connections should be identified and how they should be interpreted. This challenge may limit any meaningful interpretations of functional connectivity in the context of memory encoding and retrieval.

Our approach in addressing this challenge capitalizes on an important feature of functional connections that mediate communication in the human brain. Whether through direct or indirect synaptic connections, changes in local field potential signals coordinated across regions, or even the propagation of electric fields, communication between brain regions must occur through a physical connection of some kind [3, 15]. Importantly, the constraints associated with any physical connection imply that the time required for communication should be well defined and preserved. Every time information is conveyed from one point to another, it should take approximately the same amount of time to do so. Hence, one plausible requirement for identifying a functional connection between brain regions is that the communication between them occurs with a consistent time delay. Here, we investigate these functional connections in iEEG recordings during the paired-associates verbal memory task. Using the temporal precision afforded by iEEG, we used time-lagged mutual information (MI) to identify functional connections in the human brain that exhibit a consistent and significant increase in MI with a specific time delay. Mutual information (MI) is

attractive in that it captures all temporal relations between two brain regions [23], and is agnostic to the particular neural mechanism underlying how those regions are communicating. Each significant connection identified in this manner is therefore defined by its latency, its direction, and the consistency of its communication at that latency. We examine these identified functional connections, how these connections mediate communication during memory encoding, and how such communication is reinstated during retrieval. Our data provide novel insights into the network mechanisms that bind distributed patterns of neural activity during memory formation and retrieval.

8.1 Mutual Information Methods

To identify functional connections between electrode pairs, we first estimate the mutual information between iEEG signals captured at every electrode. We divide voltage values from each iEEG trace during one second epochs into discrete voltage bins and, for every electrode pair, estimate the joint and marginal probability distributions of voltages to calculate mutual information (MI). We repeat this for all one second epochs in a 30 second block of resting state data (98% overlap, 20 msec step size), and define the mutual information for that electrode pair as the average of all MIs from all one second epochs during that block. In order to determine whether any one pair of electrodes exhibits a significant functional connection, we then examine how the calculated MI between electrode pairs depends on the time lag between them. Within every block, we repeated our calculation of MI for every electrode pair, this time using the iEEG signal from one electrode and the iEEG signal from the other electrode captured at a different temporal delay. For every electrode pair, we therefore calculated the relation between MI and temporal delay for all temporal delay from -250 msec to 250 msec. The temporal delay that exhibits the maximum MI reflects the preferred latency and direction for that connection, and the magnitude of

MI at that preferred latency reflects how consistent that the increase in MI for that connection is at that temporal delay.

All electrode pairs exhibit some preferred latency using this approach. Importantly, however, we are interested in identifying only connections with a temporally consistent and significant increase in MI at a single preferred latency, which would suggest a significant functional connection between two brain regions with a specific time delay. Hence, for each subject we define a global threshold for the magnitude of MI at the preferred latency that corrects for multiple comparisons across all electrode pairs (Bonferroni). This threshold directly reflects how temporally consistent the communication between any pair of electrodes must be in order for their functional connectivity to be considered significant. We only consider electrode pairs with a magnitude of MI at its preferred latency that exceeds this threshold as functionally connected. Most electrode pairs exhibit an MI that does not exceed this threshold, suggesting that there is no consistent temporal relation between them, and we define those pairs as unconnected. In this manner, we only focus any analyses regarding network communication on electrode pairs that exhibit a significant functional connection with a consistent and preferred latency and direction during the resting state.

Having established the presence of electrode pairs that demonstrate significant MI at consistent latencies, we next build functional networks using only these functional connections for each participant. In our previous work, we have shown that functional networks constructed in this manner exhibit significant stability in their connections over multiple time scales. If two electrode pairs are functionally connected during one period of time, it is very likely that the same connection will be present minutes, hours, and even days later, reinforcing the suggestion that these functional connections reflect true physical connections between brain regions. It is important to note, however, that if an electrode pair exhibits a significant functional connection, the magnitude of mutual information between them is not constant across

all time. While on average, that electrode pair will exhibit an elevated level of MI (at that defined preferred latency), the level of MI between the two electrodes itself also exhibits temporal fluctuations. Indeed, an electrode pair with a true functional connection exhibits periods of time when MI is elevated and other periods when it is not, suggesting that the communication between regions effectively turns on and off. The time course of this communication in our identified functional connections forms the basis of subsequent analyses. We have previously demonstrated that distributed patterns of spectral power are indeed reinstated from memory encoding to memory retrieval. Here, we use a similar metric to assess the extent to which distributed patterns of functional connectivity are reinstated between encoding and successful retrieval. To assess the reinstatement of these communication networks, we compute MI for every identified electrode pair (using that pair's preferred latency) during sliding 1000 msec temporal windows, stepped every 100 msec, for all time points following word presentation during the encoding period, and for all time points prior to vocalization during the recall period. In this manner, across all identified electrode pairs, we construct a representation of communication for every time epoch during encoding and retrieval. Following our approach analyzing similarity in spectral power, we assess the similarity of this multidimensional representation of network communication between all encoding and recall times in order to assess whether these functional network interactions are also reinstated during successful recall.

8.2 Mutual Information Results

We performed the analysis on 9 subjects that participated in the paired associates task. The reinstatement maps averaged across subjects are shown in Figure 21. Three individual subjects are shown in Figure 22. The patterns of MI reinstatement that we see are strikingly similar to the patterns of theta power reinstatement. For each subject, the reinstatement map was computed for both theta power and MI, the two-

dimensional maps were vectorized, and the Pearson correlation was computed between them, showing a strong correlation ($R = 0.63 \pm 0.053$). This suggests that MI is closely related to theta power, and this warrants further investigation. Using the methods we described, we were able to identify stable network connections that exhibited specific time lags. If mutual information is closely related to theta power, then the task related changes in MI that we observe are likely occurring on a different timescale than the changes in MI that were used to identify the stable network connections. This is because the changes in MI that are used to identify stable networks are sensitive to time lags on the order of 10ms. The task related changes in theta and MI that we see are occurring over seconds. Therefore, it is difficult to interpret the meaning of the MI reinstatement result. There remain many open questions about the brain network connectivity during memory encoding and retrieval.

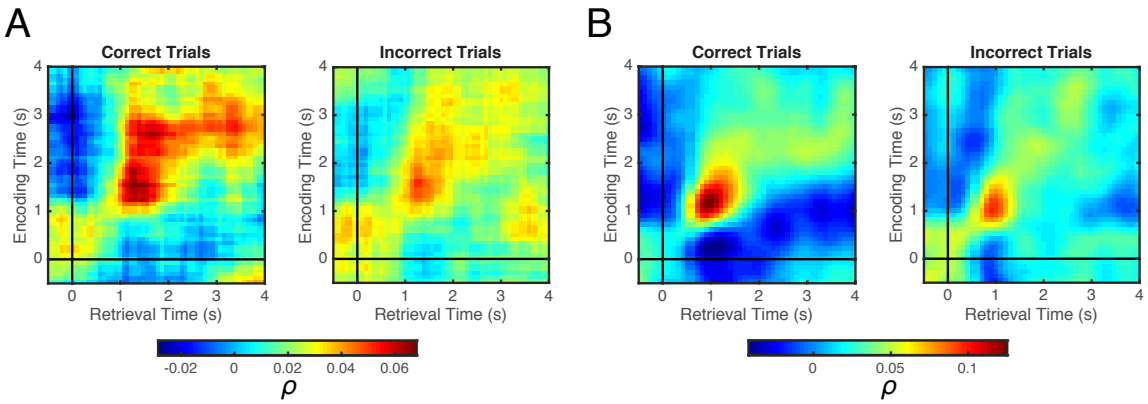

Figure 21 (A) Reinstatement map for correct and incorrect trials showing reinstatement of state vectors defined by mutual information. Time 0 corresponds to presentation of the pair/cue in encoding/retrieval. (B) Reinstatement map for correct and incorrect trials for theta power. Time axes same as (A).

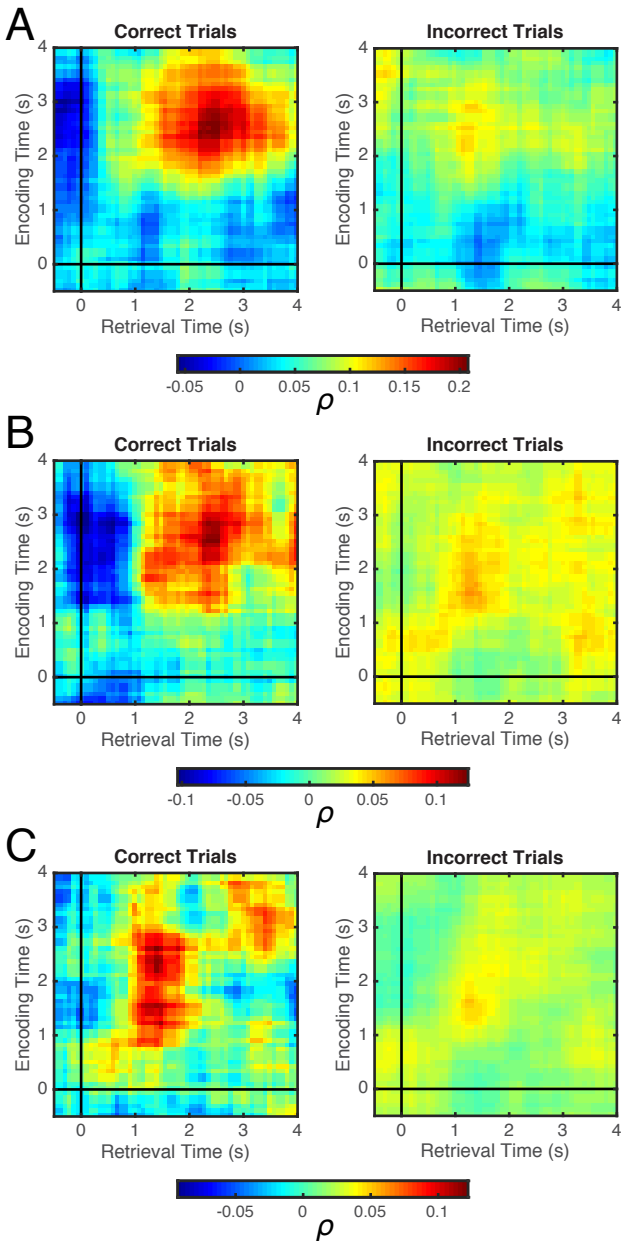

Figure 22 (A) Reinstatement map for correct and incorrect trials showing reinstatement of state vectors defined by mutual information for a single subject. Time 0 corresponds to presentation of the pair/cue in encoding/retrieval. (B and C) Same as A for additional subjects.

9 Conclusions

Our data demonstrate that reinstatement of a drifting neural representation of temporal context occurs during successful memory retrieval with precise spatiotemporal dynamics. Our approach capitalizes on the sensitivity of examining multivariate activity [37] and extends previous work in several ways. We demonstrate that patterns of spectral power distributed across multiple spatial locations and multiple frequency bands gradually change over multiple time scales. By examining trial specific reinstatement during correct and incorrect memory retrieval, we show that during successful retrieval this neural signal is recovered, providing direct evidence of mental time travel hypothesized to underlie episodic memory formation [2, 18, 38, 41, 49]. We subsequently identify the precise contributions of separate anatomic locations and frequency bands to neural reinstatement both in individuals and across our population of participants. We found that successful neural reinstatement is largely mediated by cortical high gamma activity that precedes theta activity in the temporal lobe during encoding, but this difference in timing is absent during retrieval.

One requirement of the hypothesis that successful retrieval involves mental time travel is that the neural representation of context gradually changes with time [13, 18, 41]. Consistent with this, we found that distributed patterns of spectral activity exhibited a slow temporal drift across multiple time scales. Temporal autocorrelations have traditionally been attributed to the statistical properties of complex neural signals, including resting state functional imaging [51], scalp and intracranial EEG [30]. But one possibility is that these correlations actually reflect a slowly changing representation of temporal context shaped by both external inputs and a continuously changing internal state [13, 49]. This representation should be accessed by the brain if mental time travel governs memory formation and retrieval [21, 41].

Hence, the second requirement of the mental time travel hypothesis is that retrieval recovers this gradually changing signal [18, 41, 49]. By comparing reinstatement dur-

ing correct and incorrect retrieval, our findings build upon recent evidence that trial specific reinstatement of this neural representation of temporal context occurs during successful retrieval [20, 32]. If successful retrieval is associated with the recovery of a temporal context representation, then we would expect to find a graded decrease in reinstatement as retrieval periods were paired with encoding periods separated by longer time intervals in both the forward and backward direction [18, 20, 41]. Consistent with this prediction, our data demonstrate a contiguity effect for correct trials within lists. Because the list length used here was only four items, drawing significant conclusions using only correct trial contiguity is limited and raises the possibility that the observed effects may just reflect the reinstatement of content rather than context [41]. However, we also found that incorrect trials, used here as a control condition, exhibited similarity that was principally shaped by the drifting temporal context representation itself and not by the recovery of this signal. Hence, by directly comparing reinstatement contiguity between correct and incorrect trials, our data are consistent with the hypothesis that the reinstatement of temporal context occurs during correct retrievals, and suggest that an underlying component characterized by temporal drift is evident and unmasked during incorrect trials.

Although we demonstrate a significant difference in reinstatement between correct and incorrect trials, one concern is that the observed differences in reinstatement arise as a result rather than a cause of successful retrieval [44]. While our experimental paradigm offers some control over when reinstatement and retrieval occur, our data are unable to explicitly distinguish causality between one and the other. A second possible confound in these data is that neural features responsive to visual presentation, activated both during encoding and retrieval, underlie reinstatement, and that the observed differences between correct and incorrect trials arise solely from differences in response time distributions. We note, however, that when locked to cue presentation when response time differences do not affect reinstatement, we

also observe significant differences in reinstatement (Fig. 15). We corrected for the possible confounds associated with visual responses and confirmed that a significant component of the observed reinstatement is separately mediated by memory encoding and retrieval mechanisms (Fig. 16).

We found significant differences in reinstatement when we examined distributed patterns of spectral power, but our analysis also enables us to investigate how individual spatial locations and frequencies contribute to reinstatement. We found that the largest frequency contributions to cortical reinstatement arose from theta and high gamma frequencies, consistent with the role these frequency bands play during memory encoding and retrieval [17, 46]. Within these frequency bands, we found that the greatest anatomic contributions arise from the temporal and occipital lobes, although the ventrolateral prefrontal cortex and the right temporo-parietal junction were also involved. These contributions are also not unexpected, as temporal lobe regions have been implicated in verbal memory processes, and temporal and prefrontal regions may be involved in maintaining temporal context [8, 21].

Previous studies that provide empiric support for neural reinstatement have not investigated the temporal dynamics mediating this process [7, 10, 42, 47]. We build upon these studies, as the intracranial recordings we use here enable us to examine neural reinstatement with high temporal precision across multiple frequency bands. Hence, the timing of reinstatement can inform us about the timing of encoding and retrieval processes that are particularly relevant for successful memory formation. Notably, we found that reinstatement associated with each frequency band emerges in a temporally precise manner. When both theta and high gamma frequencies were involved in reinstatement in a single location, we found that the difference in timing between encoding theta and high gamma activity was significantly greater than the difference during retrieval. These data provide evidence that memory relevant spectral power within individual cortical locations during encoding may be compressed

in time during retrieval. One possibility may be that theta and high gamma activity coordinate separate processes relevant for memory, such as attention or semantic processing, and that these processes occur separately in time during encoding and simultaneously during retrieval. However, another possibility is that the observed temporal compression complements previous studies suggesting that retrieval reactivates internal representations of an experience on a faster timescale than the original experience [6, 9]. We quantified the differences in the timescales of spectral activity during encoding and retrieval and showed that on average, both theta and high gamma power exhibit activity on a faster timescale during retrieval, which supports this hypothesis.

Together, our data demonstrate that successful retrieval is associated with reinstatement of a gradually changing temporal context representation, providing direct evidence of mental time travel. Furthermore, individual patterns of spectral power contribute to overall reinstatement and exhibit their own time course of activation during encoding and retrieval. It is notable that, given the observed temporal variability of individual frequency bands and spatial locations, averaging across all neural features still demonstrated significant differences in reinstatement between correct and incorrect trials across participants. Neural reinstatement may actually be much stronger than our analysis would suggest and it is possible that if all frequencies and spatial locations involved in encoding and retrieval were selectively chosen, neural reinstatement would be observed with higher fidelity (Fig. 17). We also present some preliminary evidence that the underlying functional connectivity networks are reinstated between encoding and retrieval. Taken as a whole, our data suggest that these distributed patterns of activity are coordinated to create a representation of temporal context which is then recovered during the retrieval of a coherent memory trace.

There are a number of ways this work may be extended. Firstly, the analysis showing a reinstatement of temporal context (Fig. 11) only considers item lags from

−2 to 2. A version of the paired associates task with more pairs per list would allow for that analysis to be extended to larger item lags in order to verify our findings. Secondly, our results indicate that there is a pair specific representation in the neural data. A task in which there are multiple presentations of the same pairs would allow us to explicitly measure the pair specific representations. Both of these projects are underway.

Ultimately, we would like to use the results from this study to build a stimulation paradigm that could improve memory performance in patients with memory disorders. This study provides insight into the patterns of spectral activity that are biomarkers of successful memory encoding and retrieval. The long term application of this study is to develop a stimulation protocol that drives patterns of spectral activity in specific brain regions at specific times to improve memory function.

References

[1] P. S. Addison. *The illustrated wavelet transform handbook: introductory theory and applications in science, engineering, medicine and finance.* Institute of Physics Publishing, Bristol, 2002.

[2] P. Alvarez and L. R. Squire. Memory consolidation and the medial temporal lobe: a simple network model. *Proceedings of the National Academy of Sciences U S A*, 91(15):7041–7045, 1994.

[3] Ed Bullmore and Olaf Sporns. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10(3):186–198, 2009.

[4] John F Burke, Nicole M Long, Kareem A Zaghloul, Ashwini D Sharan, Michael R Sperling, and Michael J Kahana. Human intracranial high-frequency activity maps episodic memory formation in space and time. *Neuroimage*, 85:834–843, 2014.

[5] John F Burke, Ashwini D Sharan, Michael R Sperling, Ashwin G Ramayya, James J Evans, M Karl Healey, Erin N Beck, Kathryn A Davis, Timothy H Lucas, and Michael J Kahana. Theta and high-frequency activity mark spontaneous recall of episodic memories. *The Journal of Neuroscience*, 34(34):11355–11365, 2014.

[6] Margaret F Carr, Shantanu P Jadhav, and Loren M Frank. Hippocampal replay in the awake state: a potential substrate for memory consolidation and retrieval. *Nature Neuroscience*, 14(2):147–153, 2011.

[7] J.F. Danker and J.R. Anderson. The ghosts of brain states past: Remembering reactivates the brain regions engaged during encoding. *Psychol Bull*, 136(1):87–102, 2010.

[8] L. Davachi, J. P. Mitchell, and A. D. Wagner. Multiple routes to memory: distinct medial temporal lobe processes build item and source memories. *Proc Natl Acad Sci U S A*, 100(4):2157 – 2162, 2003.

[9] TJ Davidson, F Kloosterman, and MA Wilson. Hippocampal replay of extended experience. *Neuron*, 63:497–507, 2009.

[10] L. Deuker, J. Olligs, J. Fell, T.A. Kranz, F. Mormann, C. Montag, M. Reuter, C.E. Elger, and N. Axmacher. Memory consolidation by replay of stimulus-specific neural activity. *J Neurosci*, 33(49):19373–19383, 2013.

[11] Kamran Diba and György Buzsáki. Forward and reverse hippocampal place-cell sequences during ripples. *Nature neuroscience*, 10(10):1241–1242, 2007.

[12] A.R. Dykstra, A.M. Chan, B.T. Quinn, R. Zepeda, C.J. Keller, J. Cormier, J.R. Madsen, E.N. Eskandar, and S.S. Cash. Individualized localization and cortical surface-based registration of intracranial electrodes. *NeuroImage*, 59(4):3563–3570, 2011.

[13] W. K. Estes. Statistical theory of spontaneous recovery and regression. *Psychological Review*, 62:145–154, 1955.

[14] David J Foster and Matthew A Wilson. Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature*, 440(7084):680–683, 2006.

[15] Michael D Fox and Marcus E Raichle. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature Reviews Neuroscience*, 8(9):700–711, 2007.

[16] Christopher R. Genovese, Nicole A. Lazar, and Thomas E. Nichols. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. *NeuroImage*, 15:870–878, 2002.

[17] S. Hanslmayr and T. Staudigl. How brain oscillations form memories- a processing based perspective on oscillatory subsequent memory effects. *NeuroImage*, 121(5), June 2013.

[18] M. W. Howard, M. S. Fotedar, A. V. Datey, and M. E. Hasselmo. The temporal context model in spatial navigation and relational learning: Toward a common explanation of medial temporal lobe function across domains. *Psychological Review*, 112(1):75–116, 2005.

[19] M. W. Howard and M. J. Kahana. A distributed representation of temporal context. *Journal of Mathematical Psychology*, 46:269–299, 2002.

[20] M. W. Howard, I. V. Viskontas, K. H. Shankar, and I. Fried. Ensembles of human MTL neurons "jump back in time" in response to a repeated stimulus. *Hippocampus*, 22:1833–1847, 2012.

[21] L. J. Jenkins and C. Ranganath. Prefrontal and medial temporal lobe activity at encoding predicts temporal context memory. *Journal of Neuroscience*, 30(46):15558–15565, 2010.

[22] Ole Jensen, Jochen Kaiser, and Jean-Philippe Lachaux. Human gamma-frequency oscillations associated with attention and memory. *Trends in neurosciences*, 30(7):317–324, 2007.

[23] Jaeseung Jeong, John C Gore, and Bradley S Peterson. Mutual information analysis of the eeg in patients with alzheimer's disease. *Clinical Neurophysiology*, 112(5):827–835, 2001.

[24] J.D. Johnson, S.G.R. McDuff, M. D. Rugg, and K.A. Norman. Recollection, familiarity, and cortical reinstatement: A multivoxel pattern analysis. *Neuron*, 63:697–708, 2009.

[25] J.D. Johnson and M.D. Rugg. Recollection and the reinstatement of encoding-related cortical activity. *Cerebral Cortex*, 2007.

[26] Michael J Jutras and Elizabeth A Buffalo. Synchronous neural activity and memory formation. *Current opinion in neurobiology*, 20(2):150–155, 2010.

[27] M. J. Kahana. *Foundations of Human Memory*. Oxford University Press, New York, NY, 2012.

[28] M. J. Kahana, M. W. Howard, and S. M. Polyn. Associative retrieval processes in episodic memory. In H. L. Roediger, III, editor, *Cognitive psychology of memory. Vol. 2 of Learning and memory: A comprehensive reference, 4 vols. (J. Byrne, Editor)*. Elsevier, Oxford, 2008.

[29] B.A. Kuhl, J. Rissman, and A.D. Wagner. Multi-voxel patterns of visual category representation during episodic encoding are predictive of subsequent memory. *Neuropsychologia*, 2011.

[30] K. Linkenkaer-Hansen, V. V. Nikouline, J. M. Palva, and R. J. Ilmoniemi. Long-range temporal correlations and scaling behavior in human brain oscillations. *Journal of Neuroscience*, 21(4):1370–1377, 2001.

[31] Joseph A Maldjian, Paul J Laurienti, Robert A Kraft, and Jonathan H Burdette. An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *Neuroimage*, 19(3):1233–1239, Jul 2003.

[32] J. R. Manning, S. M. Polyn, G. Baltuch, B. Litt, and M. J. Kahana. Oscillatory patterns in temporal lobe reveal context reinstatement during memory search. *Proc Natl Acad Sci U S A*, 108(31):12893 – 12897, 2011.

[33] Eric Maris and Robert Oostenveld. Nonparametric statistical testing of EEG- and MEG-data. *Journal of Neuroscience Methods*, 164:177–190, 2007.

[34] J. L. McClelland, B. L. McNaughton, and R. C. O'Reilly. Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychological Review*, 102(3):419–57, 1995.

[35] S.G. McDuff, H.C. Frankel, and K. A. Norman. Multivoxel pattern analysis reveals increased memory targeting and reduced used of retrieved details during single-agenda source monitoring. *J Neurosci*, 29:508–516, 2009.

[36] J. F. Miller, M. Neufang, A. Solway, A. Brandt, M. Trippel, I. Mader, S. Hefft, M. Merkow, S. M. Polyn, J. Jacobs, M. J. Kahana, and A. Schulze-Bonhage. Neural activity in human hippocampal formation reveals the spatial context of retrieved memories. *Science*, 342(6162):1111–1114, 2013.

[37] K. A. Norman, E. Newman, G. Detre, and S. M. Polyn. How inhibitory oscillations can train neural networks and punish competitors. *Neural Computation*, 18:1577–1610, 2006.

[38] K. A. Norman and R. C. O'Reilly. Modeling hippocampal and neocortical contributions to recognition memory: A complementary learning systems approach. *Psychological Review*, 110:611–646, 2003.

[39] P. L. Nunez and R. Srinivasan. *Electric Fields of the Brain*. Oxford University Press, New York, 2006.

[40] S. M. Polyn, V. S. Natu, J. D. Cohen, and K. A. Norman. Category-specific cortical activity precedes retrieval during memory search. *Science*, 310:1963–1966, 2005.

[41] S. M. Polyn, K. A. Norman, and M. J. Kahana. A context maintenance and retrieval model of organizational processes in free recall. *Psychological Review*, 116(1):129–156, 2009.

[42] J. Rissman and A. D. Wagner. Distributed representations in memory: Insights from functional brain imaging. *Annual Review of Psychology*, 63:101–128, 2012.

[43] M. Ritchey, E.A. Wing, K.S. LaBar, and R. Cabeza. Neural similarity between encoding and retrieval is related to memory via hippocampal interactions. *Cerebral Cortex*, 23(12):2818–2828, 2013.

[44] M.D. Rugg, J.D. Johnson, H. Park, and M.R. Uncapher. Encoding-retrieval overlap in human episodic memory: a functional neuroimaging perspective. *Progress in Brain Research*, 169:339–352, 2008.

[45] P. B. Sederberg, M. W. Howard, and M. J. Kahana. A context-based theory of recency and contiguity in free recall. *Psychological Review*, 115(4):893–912, 2008.

[46] P. B. Sederberg, A. Schulze-Bonhage, J. R. Madsen, E. B. Bromfield, D. C. McCarthy, A. Brandt, M. S. Tully, and M. J. Kahana. Hippocampal and neocortical gamma oscillations predict memory formation in humans. *Cerebral Cortex*, 17(5):1190–1196, 2007.

[47] B.P. Staresina, R. N. Henson, N. Kriegeskorte, and A. Alink. Episodic reinstatement in the medial temporal lobe. *J Neurosci*, 32(50):18150–18156, 2012.

[48] Nanthia Suthana, Zulfi Haneef, John Stern, Roy Mukamel, Eric Behnke, Barbara Knowlton, and Itzhak Fried. Memory enhancement and deep-brain stimulation of the entorhinal area. *New England Journal of Medicine*, 366(6):502–510, 2012.

[49] E. Tulving. Episodic and semantic memory. In E. Tulving and W. Donaldson, editors, *Organization of Memory.*, pages 381–403. Adademic Press, New York, 1972.

[50] Sabine Weiss and Peter Rappelsberger. Long-range eeg synchronization during word encoding correlates with successful memory performance. *Cognitive Brain Research*, 9(3):299–312, 2000.

[51] E. Zarahn, G. K. Aguirre, and M. D'Esposito. Empirical analysis of bold fmri statistics. *NeuroImage*, 5(3):179–197, April 1997.

Author Information

Robert Benjamin Yaffe
Born: May 4, 1989 in Washington, D.C., USA

B.A.S. in Computer Science from the University of Pennsylvania, 2011

Ph.D. in Biomedical Engineering from the Johns Hopkins University, 2016
Graduate Partnerships Program, National Institute of Neurological Disorders and Stroke, 2013 - 2016
Mentors: Sridevi Sarma, Ph.D. and Kareem Zaghloul, MD, Ph.D.

Teaching Experience:
Fall 2014, JHU: Melding Mind and Machine: Exploring neurotechnologies that can read our minds and alter our brains. Developed and taught course.
Fall 2014, JHU: Principles of Biomedical Engineering Instrumentation, TA
Spring 2014, JHU: Biological Models and Simulations, TA
Spring 2013, JHU: Applied Neuroscience and Neurotechnology. Developed and taught course.